INVESTIGATION OF THE START BACK SCREENING TOOL IN OUTPATIENT PHYSICAL THERAPY SETTINGS

By

JASON MICHAEL BENECIUK

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

Dedicated to Monika, Aiden, Alexa, Brody & Barren for being my inspiration and putting up with my “mood-swings” over the past five years; you guys have taken this journey with me – and for that I thank you

ACKNOWLEDGMENTS

I would like to thank all of those people who helped make this dissertation possible. First and most important, my committee chair, Dr. Steven George for his commitment, guidance, and dedication over the past five years. I'm not sure he knew what he was getting himself into when my wife came to the application interview and had more questions than I did; but for whatever reason he still agreed to take me under his wing and for that I am truly grateful. Second, I would like to thank my committee members for their contributions in my overall professional development. Dr. Mark Bishop has always reminded me to stay focused on the big picture, especially when I had a tendency to get 'lost in translation'. Fortunately, Dr. Michael Robinson has managed to sit through my repeated 'change analysis' questions and remained on my committee – although he probably just felt sorry for me. Dr. Nabih Asal has provided me with an epidemiological perspective on this project and served as my primary public health mentor. Dr. Julie Fritz has provided me with insight related to the classification of low back pain and has offered me the opportunity to be involved with a similar project to this dissertation in the past; therefore I am thankful that she has agreed to be a member of this dissertation committee. Third, I would like to thank Brooks Rehabilitation for funding this project. Fourth, I would like to acknowledge the T-32 Neuromuscular Plasticity Training Grant and Dr. Krista Vandenborne for funding my pre-doctoral education. Finally, thank you to Joel Bialosky, Carolina Valencia, Roy Coronado, and Corey Simon for their support and feedback over the years. A special thank you to Anne Nisenzon, Laura Wandner, and Calia Torres is deserved for their roles in this project. Last, but definitely not least, thank you to Jim Viti for being a great mentor and friend.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FABQ-PA	Fear-Avoidance Beliefs Questionnaire – physical activity scale
FABQ-W	Fear-Avoidance Beliefs Questionnaire – work scale
FAM	Fear-Avoidance Model of Musculoskeletal Pain
LBP	Low Back Pain
NPRS	Numerical Pain Rating Scale
ODQ	Revised Oswestry Disability Questionnaire
PCS	Pain Catastrophizing Scale
PII	Physical Impairment Index
PHQ-9	Patient Health Questionnaire – 9-item version
RMDQ	Roland-Morris Disability Questionnaire
STAI-T	State-Trait Anxiety Inventory – trait portion
STarT	Subgroups for Targeted Treatment Back Screening Tool
TSK-11	Tampa Scale for Kinesiophobia – 11-item version

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

INVESTIGATION OF THE START BACK SCREENING TOOL IN OUTPATIENT PHYSICAL THERAPY SETTINGS

By

Jason Michael Beneciuk

December 2011

Chair: Steven Z. George
Major: Rehabilitation Science

Low back pain (LBP) is a major public health problem representing a significant portion of patients seen in physical therapy (PT) settings where some develop chronic symptoms influenced by psychological factors. The Fear-Avoidance Model of Musculoskeletal Pain consists of individual psychological factors that provide a theoretical explanation as to why chronic pain conditions develop in a minority of those experiencing an acute episode. Currently, there is no standardized psychological screening process to identify these at-risk patients suited for clinicians in busy PT settings. One potential screening method involves measuring the influence of individual psychological constructs with several questionnaires, while an alternative method includes screening for general psychological distress with a single questionnaire. The STarT Back is a 9-item screening tool consisting of functional and psychological items that allocates patients into subgroups (i.e., low, medium, or high risk) describing risk status for future disability and are associated with initial treatment options in primary care settings. The STarT Back was developed and intended for use in primary care settings; however has potential for use in physical therapy settings.

This study included 146 patients seeking PT for LBP that completed a battery of measures at initial evaluation and 4-weeks later. Results indicated that the STarT Back was consistent at differentiating patients based on severity of LBP at initial evaluation but not as consistent for predicting 4-week treatment outcomes. Furthermore, an individual Fear-Avoidance Model psychological measure (FABQ-PA) was useful in predicting which patients maintained STarT Back high risk status following 4-weeks of PT treatment for LBP.

The primary limitation to this study was that PT treatment was not tailored to address LBP associated psychological factors. This study provides preliminary evidence that a brief psychological screening tool maybe useful in distinguishing severity of LBP, but may have limited ability to predict 4-week treatment outcomes.

CHAPTER 1
INTRODUCTION

From a public health perspective, low back pain (LBP) is the second most common cause of adult disability in the United States.[1, 2] Pain has been recognized by the World Health Organization as a problem of epidemic proportion,[3] with others highlighting the influence of LBP specifically.[4, 5] This has important implications for rehabilitation science considering one of the largest proportions of direct medical costs for LBP have been attributed to physical therapy services.[6] As a result, the need for effective interventions has been emphasized with particular concern for the influence that modifiable factors have on LBP outcomes,[7-9] with psychological factors specifically being highlighted as a potential treatment target.[5, 10-13] Therefore, one goal for rehabilitation scientists is to determine optimal screening procedures for patients with LBP in order to identify those at increased risk for developing chronic symptoms.

In 2009, the Tenth International Forum for Primary Care Research on Low Back Pain was held in which recent concepts, research methods, and relevant study results were presented.[5] Consistent with the general theme of this dissertation, “early risk factor screening” for poor clinical outcomes was identified as a potential method to improve the efficiency and effectiveness of care and associated clinical outcomes.[5] Ultimately, an optimal screening process would provide clinicians with valuable information in the form of prognostic indicators and/or treatment effect modifiers that would assist in clinical decision-making.

Specific to physical therapy, “psychologically informed practice” has recently been presented as a secondary prevention approach for chronic LBP that integrates both biomedical (focused on physical impairments) and cognitive-behavioral (focused on

psychological distress) principles.[13] The primary goal of a psychologically informed approach is to prevent future LBP-associated disability with routine and specific identification of modifiable psychological risk factors being emphasized. Specific to this dissertation, determining the validity and clinical utility of commonly used psychological screening measures has been indicated as a top priority for future research for psychologically informed practice.[13]

Therefore, the primary goals of this dissertation were to: 1) determine the validity of the STarT Back Tool (a multi-construct psychological screening measure) in outpatient physical therapy settings and 2) test the clinical utility of the STarT Back Tool in comparison to commonly used single-construct psychological screening measures. The following literature review will provide relevant information to support the rationale for these goals.

CHAPTER 2
LITERATURE REVIEW

Epidemiology of Low Back Pain

Incidence and Prevalence

An overlying goal of epidemiological research includes the control of health problems via the study of the distribution and determinants of health-related conditions.[14] The heterogeneity associated with studying pain in human populations complicates the ability to compare findings across epidemiological studies. This is particularly relevant when studying certain types of musculoskeletal pain, such LBP where psychological factors have an influential role in the pain experience.[10] Considering this heterogeneity, it is important to critically evaluate the case definition used to define LBP when interpreting the results of epidemiological studies and the impact LBP has on society prior to implementing results into decision-making processes related to rehabilitation.

Commonly reported epidemiological occurrence rates include incidence and prevalence. Incidence has traditionally been used to indicate the frequency of a new developed case of a particular health-related outcome.[14] Prevalence measures the frequency of an existing outcome based on specific time parameters (i.e., point, period, or lifetime prevalence).[14] In terms of epidemiological studies involving LBP, prevalence estimates are frequently the reported measure of occurrence based on the episodic or recurrent nature of this condition; thereby limiting the ability to report true incidence rates. It is important to note that incidence can also refer to the occurrence of a 'new' or 'recent' episode of LBP, although it may not be the initial incidence of LBP over the course of one's lifetime. However, variable methods in defining episodes or recurrence

have been used in review studies,[15] suggesting a need for standardized LBP recurrence terminology.[16] For example, in a previous study, 73% of patients with acute LBP had at least one reoccurrence within 12-months,[17] however when using standardized methods to define recurrence, estimates have been reported to be much lower (i.e., 24% to 33%).[18] Considering these factors, LBP occurrence rates will be reported as prevalence estimates in this dissertation as they are most commonly reported.

The prevalence of LBP is extensively documented in the literature; however estimates vary due to differences in research methodology. For example, case definitions used to describe the characteristics and duration of LBP in the general population are inconsistent, thus limiting the ability to compare prevalence estimates among studies. A systematic review[19] of the literature from 1966 to 1998 consisting of population prevalence studies has provided very broad point prevalence (12% to 33%), 1-year prevalence (22% to 65%), and lifetime prevalence (11% to 84%) estimates of LBP. Although direct comparisons cannot be made, findings from a separate systematic review highlight the influence that case definitions have on prevalence estimates. Loney and Stratford[20] provided point prevalence (4% to 33%) and 1-year prevalence (4% to 63%) estimates of LBP from prevalence studies consisting of adult community-based samples but excluded studies involving occupational groups (e.g., healthcare works, industrial workers, military). The authors commented that improved methods are required prior to using these estimates with confidence when considering issues related to healthcare policy (e.g., physical therapy), including the standardization of case definitions for adequate comparisons among studies.[20]

The National Health Interview Survey (NHIS) is a national survey conducted by the National Center for Health Statistics, a branch of the Centers for Disease Control and Prevention. The NHIS provides U.S. national estimates for a broad range of health measures among the civilian, non-institutionalized adult population. Deyo and colleagues[21] provided age-adjusted prevalence estimates of LBP using 2002 NHIS data (n = 31,044), in which 26.4% of respondents indicated they experienced LBP "lasting at least a whole day in the past 3 months". Additional findings indicate that: 1) adults greater than 45 years of age reported LBP at an elevated rate in comparison to younger adults (18 to 44 years), however prevalence decreased slightly among the oldest adults (≥ 75 years) and 2) women (28.3%) reported a higher frequency of LBP in comparison to men (24.3%).[21] Strine and Hootman[22] also utilized 2002 NHIS data (n = 29,828) to provide prevalence estimates of LBP. Based on weighted data to represent national estimates, 17.0% reported LBP, resulting in an estimated 34 million adults in the U.S. population. Potential reasons for differences in prevalence estimates (i.e., 26.4% vs. 17.0%) utilizing 2002 NHIS data may be a result of different weighting techniques, method of estimate reporting, and / or sample sizes.

The National Health and Nutrition Examination Survey (NHANES) is another national survey conducted by the National Center for Health Statistics, a branch of the Centers for Disease Control and Prevention. NHANES is a program of studies designed to assess the health and nutritional status of non-institutionalized adults and children in the United States and is unique in that it combines interviews and physical examinations. Hardt and colleagues[23] provided estimates of LBP prevalence using 1999 to 2002 NHANES data (n = 10,291), in which 10.1% of respondents indicated they

experienced “non-minor LBP, ≥ 3 months in duration, during the previous month, lasting ≥ 24 hours”.

Freburger and colleagues[24] reported on the rising prevalence of chronic and acute LBP following a cross-sectional evaluation of North Carolina households conducted in 1992 and repeated in 2006. Chronic LBP was defined as: 1) LBP that limited daily activity everyday for the past 3 months or 2) > 24 episodes of LBP that limited activity ≥ 1 day in the past year. Provided estimates indicate an increasing prevalence of chronic, impairing LBP over the 14-year interval from 1992 (3.9%) to 2006 (10.2%), resulting in an overall increased prevalence of 162%. Increases were identified in all age strata, in men and women, and in White and Black races.

In the Freburger and colleagues study,[24] acute LBP was defined as: 1) LBP limiting daily activity ≥ 1 day and < 3 months or 2) < 25 episodes of LBP that limited activity in the past year. Provided estimates indicate an increasing prevalence of acute LBP over the 14-year interval from 1992 (7.3%) to 2006 (10.5%), resulting in an overall increased prevalence of 44%.[24] The authors speculated on the increased prevalence of several factors which may have been associated with increased LBP prevalence rates in North Carolina during this interval including obesity, psychosocial and physical work demands, depression, and increased symptom awareness and reporting of LBP. While these results cannot be extrapolated to national estimates, this study is relevant because it is the first population-based study in the United States which has examined trends in the prevalence of LBP using identical case definitions of LBP.

Focus on Therapeutic Outcomes (FOTO), Inc (Knoxville, Tennessee) is an international medical rehabilitation database.[25] Between 2002 and 2006 data was

analyzed from 17,804 patients being treated for neuromusculoskeletal conditions in 121 outpatient rehabilitation clinics in 26 states (in the United States). Of these patients, 31.4% received outpatient physical therapy services for conditions related to musculoskeletal LBP.[26] These findings are similar to those found between 2005 and 2008 in Israel outpatient rehabilitation clinics.[27]

In summary, prevalence estimates for LBP in the U.S. have ranged from 17% to 26% and chronic LBP rates specifically have been estimated at 10%.[21-23] North Carolina data indicates that the prevalence of acute and chronic LBP has increased 44% and 162% respectively, when standard definitions were used over a 14-year period.[24] Specific to outpatient physical therapy settings, data has indicated that 31% of patients seeking outpatient physical therapy seek services for LBP.[26]

Societal Impact

Considering symptoms of pain are a leading reason for medical visits[28, 29] and the most common pain complaints are musculoskeletal (of which LBP is the most common),[30] it is plausible to assume that healthcare expenditures associated with LBP have an enormous negative impact from a societal perspective. In 2005, the Centers for Disease Control and Prevention indicated that “back or spine” problems were the second leading cause of disability in the United States.[2] In 2002, LBP accounted for approximately 2.3% of all physician visit rates with only routine examinations, hypertension, and diabetes resulting in more visits.[21] In 2008, back pain was the fourth leading reason for hospital outpatient or office based provider visits.[31] Results from the 2007 Medical Expenditure Panel Survey indicated that 27 million adults (11.9%) reported LBP of which 70.4% received treatment totaling 30.3 billion dollars amounting in annual mean expenditures of approximately 1,600 dollars per adult.[32] In 2005, direct

treatment for back and neck pain accounted for approximately 86 billion dollars in healthcare expenditure in the United States, however did not distinguish expenditures associated with LBP specifically.[33] Additionally, a systematic review of LBP cost of illness studies in the United States and internationally between 1997 and 2007, indicated that the largest proportion of direct medical costs for LBP can be attributed to physical therapy (17%) and inpatient services (17%), followed by pharmacy (13%) and primary care (13%).[6] In that systematic review, the range of direct (12.2 to 90.6 billion dollars) and indirect (7.4 to 28.2 billion dollars) costs of LBP in the United States were determined from the results of 3 and 5 studies respectively. Special consideration is warranted when interpreting the upper-limit of direct costs, as that review study was published in 1998 and results may be a reflection of the 1996 approval of the fusion cage surgical implant. The authors commented that the most notable finding in their review of LBP cost of illness was the heterogeneity in methodology used to derive cost of illness among 27 studies examined.

While the above study findings demonstrate the burden that LBP has on society in the form of healthcare expenditures, they do not provide information on the entire healthcare experience associated with an episode of LBP. A previous study found that individuals with back pain incurred healthcare expenditures 60% higher than individuals not experiencing back pain, however did not account for individual differences in those with back pain.[34] Studies investigating the entire healthcare experience throughout a LBP episode may provide further insight as to how LBP may also be associated with non-LBP costs. Nimgade and colleagues[35] found that average monthly healthcare expenditures for non-LBP expenditures increased when compared to the previous 1 to 3

months following the initiation of a LBP episode. Moreover, patients with increased non-LBP expenditures were more likely to have been prescribed psychiatric medications.[35]

LBP has been reported to be a common reason for lost work days,[36-38] work-related disability,[36] and the fourth most costly physical health condition affecting several large U.S. employers.[39] It has been estimated that 149 million days of work are lost because of LBP.[40] In a random sample of U.S. working adults 18 to 65 years of age over a two week period, LBP ranked second among painful conditions as the reason for lost productive time at work (mean = 5.2 hours per week), costing employers an estimated 19.8 billion dollars per year.[38] In the United States, working adults 40 to 65 years of age with LBP cost employers an estimated 7.4 billion dollars per year, with those experiencing an exacerbation of LBP symptoms accounting for 71.6% of this cost.[37] These findings may be underestimating actual costs based on the results of a systematic review indicating that when total costs are reported, the indirect costs resulting from lost work productivity represent a majority of overall costs associated with LBP.[6]

In summary, direct costs associated with healthcare services for LBP are enormous and there have been indications that physical therapy services account for a large portion of these costs. Furthermore, indirect costs are frequently associated with greater adverse effects on society with the development and progression to chronic LBP being highlighted.

Healthcare Utilization for Low Back Pain: Who is Seeking Care?

Results from a 1997 North Carolina survey[41] indicated that 12.6% of respondents sought physical therapy services for acute LBP with post high school education,

receiving workers compensation, having had prior physical therapy for LBP, having LBP and pain below the knee, and increased disability scores being positively associated with physical therapy utilization. Results from a 2011 North Carolina population-based survey[42] indicated that 29.7% of respondents sought physical therapy services for chronic LBP with receiving workers compensation, having seen a physician specialist, and higher levels of function being positively associated and having no health insurance being negatively associated with physical therapy utilization. Furthermore, in that study, Freburger and colleagues[42] found that some previously reported effective physical therapy treatments for chronic LBP were underutilized (e.g., spinal manipulation – 10.4%) and ineffective treatments were overutilized (e.g., corset or bracing – 24.0%); potentially indicating the need to improve processes of matching patients with optimal treatment approaches for LBP.

A dose repose relationship has also been indicated between the timing of physical therapy initiation for acute LBP and subsequent healthcare utilization. Gellhorn and colleagues[43] analyzed data from a nationally representative sample of Centers for Medicare and Medicaid Services physician outpatient billing claims and found that patients (mean age = 76.0 years) who received physical therapy early after an episode of acute LBP were at lower risk for subsequent LBP-related healthcare utilization over the following year compared to those who received physical therapy at later times. Furthermore, only 16.2% of patients received physical therapy for LBP, potentially suggesting underutilization of physical therapy services for patients with acute or subacute LBP by medical generalist specialists.[43]

Collectively, these findings potentially suggest that the majority of individuals experiencing LBP are not necessarily utilizing the bulk healthcare resources – especially physical therapy services, which Freburger and colleagues have referred to as "missed opportunities".[42] Rather, a small percentage of patients with chronic LBP account for a large fraction of associated costs.[44, 45] Furthermore, the timing of physical therapy initiation during early episodes of LBP and improvements in processes that aim to match patients with optimal treatment approaches are potentially important factors to consider when attempting to lower subsequent healthcare utilization for LBP.

The Fear-Avoidance Model of Musculoskeletal Pain

The Fear-Avoidance Model of Musculoskeletal Pain (FAM) provides a theoretical explanation as to why chronic pain conditions develop in a minority of those experiencing an acute episode. Considering the experience of pain is highly variable among individuals, an understanding of what constitutes 'pain perception' is an important concept to comprehend in order to gain a full appreciation of the FAM. As proposed by Lethem and colleagues,[46] pain perception involves a 'sensory component' (i.e., pain sensation) and an 'emotional reaction component' (i.e., pain experience, pain behavior, physiological responses to pain stimulation). Potential coping strategies used by individuals experiencing pain range from the extremes of 'confrontation' to 'avoidance'. The type of coping strategy adopted is influenced by several psychosocial factors (i.e., stressful life events, personal pain history, personal coping / response strategies, and personality).[46] Adaptive responders (i.e., confronters) view pain as temporary and have increased motivation to return to full function. These individuals maintain a timely balance (i.e., synchrony) between pain sensation, pain experience, and pain behaviors thereby increasing the likelihood of resuming normal function.

Alternatively, non-adaptive responders (i.e., avoiders) are not able to maintain this timely balance (i.e., desynchrony) and may exhibit an increased fear of pain, resulting in avoidance of physical and social activities. Although the FAM is primarily focused on the emotional reaction component of pain perception, it provides a plausible explanation of how the psychological consequences of avoidance behaviors can ultimately lead to adverse physical and psychological consequences, including physical disability, depression, and exaggerated pain perception.[46]

Since its development in 1983, the FAM has gained popularity in research focusing on musculoskeletal pain and has been expanded upon. For example, Vlaeyen and colleagues[47] refined the FAM in 1995 by incorporating the term 'pain catastrophizing' into the model as a potential precursor to pain-related fear. Pain catastrophizing involves an exaggerated negative interpretation of actual and anticipated pain experience resulting in exaggerated threat value of pain and negative appraisal of one's ability to cope with pain.[48-51] Vlaeyen and Linton further refined the FAM in 2000 by speculating on the role that 'negative affectivity' (i.e., subjective distress, negative mood states)[52, 53] and 'threatening illness information' have on pain catastrophizing.[54]

The current FAM[55] accounts for several components which were previously discussed (e.g., adoption of coping strategies), however also consists of several revisions that provide additional insight to psychological constructs within the FAM. 'Fear of pain' and 'pain anxiety' are distinguished from one another, with the former representing defensive behaviors (e.g., escape) in response to present threats and the latter representing preventative behaviors (e.g., avoidance) in response to anticipated

future threats. Although both types of behaviors may be advantageous during acute pain episodes, they may influence the threshold of future pain experiences.[55] Although individual psychological constructs within the FAM are separate from one another, it is debatable as to whether they are actually distinct enough from one another to be distinguished from a clinical perspective. Nevertheless, the current FAM provides an in-depth theoretical perspective as to why some patients that experience an acute episode of musculoskeletal pain eventually develop chronic musculoskeletal pain.

Recently, a collaborative narrative review presented potential limitations and proposed potential extensions that can potentially be incorporated into the current FAM in order to increase its clinical utility.[56] For example, evidence for the current FAM is predominantly supported through experimental studies, while observational studies consisting of patients are associated with conflicting results and frequently utilize cross-sectional study designs. Furthermore, suggestions for future research to improve the clinical utility of the FAM included: 1) the development of more psychometrically sound assessment tools, 2) the ability to identify subgroups within the FAM, and 3) the need for more prospective studies.[56]

Prognosis and Low Back Pain

Methodological Issues

The ability to generalize findings from previous LBP prognostic studies is associated with limitations based on methodological issues.[57] Potential explanations as to why inconsistent findings exist in the literature include heterogeneity of populations (e.g., general population, clinical population), settings (e.g., primary care, secondary care, occupational), stage of condition (e.g., new episode, acute, subacute, chronic), and potential prognostic indicators (e.g., demographic, clinical, psychological).

Furthermore, inconsistent measurement of clinical outcome domains across studies is common; therefore the ability to generalize findings is potentially limited. Many review studies consist of individual studies that report on mixed outcome domains (e.g., return-to-work status, pain intensity, disability) which need to be interpreted with caution when attempting to discern results. Specifically, changes in different LBP outcome domains are not always positively correlated,[5, 58] which limits the ability to draw "definitive" conclusions based on review study results that incorporate mixed outcome domains.

The influence of psychological factors on LBP outcomes also varies based on screening methods (e.g., single-construct or multiple-construct approaches), and statistical analyses. Specific to interpretation of study findings and statistical analyses, it is important to distinguish between prognostic factors and treatment effect modifiers. Prognostic factors are characteristics that identify patients who recover at different rates or have different outcomes, which can be used to predict patient outcomes.[59, 60] Alternatively, treatment effect modifiers are characteristics that identify subgroups of patients who respond differently to a specific intervention (or intervention approach) that are capable of predicting treatment effects.[59, 60] While treatment effect modifiers and prognostic factors may share similar characteristics, they provide different information to both researchers and clinicians, therefore should be interpreted and analyzed separately. As a result, it has been suggested that when attempting to identify either prognostic factors or treatment effect modifiers, appropriate study designs and analysis techniques should be implemented.[59, 60] For example, if the intent is to identify prognostic factors, single arm cohort studies similar to the design used in this study are appropriate. Alternatively, if the intent is to identify treatment effect modifiers, two-arm

trials that incorporate a comparison group (preferably a control group) should be utilized to investigate factors that may modify treatment effects.

In summary, methodological inconsistencies across prognostic studies related to LBP is common, therefore establishing definitive findings to outpatient physical therapy settings is difficult. As a result, the subsequent sections can be interpreted as a brief summary of the collective evidence for the relationship of psychological factors with: 1) a new onset of LBP; 2) baseline clinical characteristics; and 3) the influence on future clinical outcomes.

New Onset of Low Back Pain

A complicating factor involving the role of psychological factors on LBP includes the "chicken or the egg" dilemma.[61, 62] Specifically, are psychological characteristics pre-existing factors that influence patients prior to an episode of LBP or do they manifest their influence following the initiation of a LBP experience? To succinctly investigate this question, a longitudinal, prospective cohort study design consisting of people without LBP would be optimal to determine patient's psychological status prior to LBP has an influence on a new episode of LBP. For example, in a classic study by Bigos and colleagues,[63] premorbid factors that were predictive of future LBP episodes were investigated in over 3000 aircraft workers. Results indicated that work perceptions and certain psychosocial responses were most predictive of subsequent reports of LBP.[63] Pain-related fear and pain catastrophizing have been implicated as risk factors associated with the development of future LBP in healthy individuals.[64-66] Other studies have suggested similar results in the general public or work settings.[62, 64, 65, 67-70]

In summary, there is evidence to suggest that elevated levels of premorbid psychological factors are predictive of new LBP episodes, which is consistent with the results of a systematic review on this topic published in 2000.[61]

Baseline Clinical Characteristics

Psychological factors have been found to be associated with baseline clinical characteristics related to LBP in cross-sectional studies. Measures of pain catastrophizing,[71-73] pain-related fear,[71, 74] fear-avoidance beliefs,[74, 75] depression,[73] anxiety,[76, 77] and negative affect[75] have all been associated with baseline clinical characteristics (e.g., pain intensity, disability). For example, elevated levels of fear-avoidance beliefs about physical activity have been found to be positively associated with baseline self-reported disability in both acute and chronic LBP patients.[75] Other studies have indicated similar relationships between fear-avoidance beliefs about work,[73] pain-related fear,[71, 74] or pain catastrophizing[71, 72] with baseline self-reported disability or pain intensity.[71-73]

Previous studies conducted by our group in outpatient physical therapy settings have reported on the relationship between individual FAM measures and baseline clinical characteristics in patients with LBP.[78, 79] Testing the criterion validity of several FAM measures (FABQ-PA, FABQ-W, FPQ-9, TSK-11) indicated that fear-avoidance beliefs about physical activity (FABQ-PA) scores contributed additional variance to initial pain intensity ratings (23%) and disability scores (23%), while fear-avoidance beliefs about work (FABQ-W) scores contributed additional variance to initial physical impairment (13%) and disability scores (8%) in patients with chronic LBP.[78] Pain catastrophizing (PCS) scores also contributed additional variance (37%) to initial depression scores. Other FAM measures did not contribute to the respective

regression models. In a separate study with a different sample of patients, testing the criterion validity of similar FAM measures (FABQ-PA, FABQ-W, TSK-11, PCS) indicated that FABQ-PA scores contributed additional variance to initial pain intensity ratings (18%) and disability scores (27%), while PCS scores contributed additional variance to initial pain intensity ratings (6%) and disability scores (3%) in patients with LBP of various durations.[79] Other FAM measures did not contribute to the respective regression models. Further analysis indicated that PCS scores mediated the relationship of the FABQ-PA by weakening its association with pain intensity ratings and disability scores.

In summary, there is evidence to support that elevated levels of psychological distress are associated with higher pain intensity, disability, and physical impairment at baseline in a variety of healthcare settings, including physical therapy.

Future Clinical Outcomes

Although previous LBP history and pain intensity may be consistent predictors of future disability,[80] psychological factors have also been suggested to be influential in the development of chronic LBP and related disability. Pain catastrophizing,[67] pain-related fear,[67, 81, 82] fear-avoidance beliefs,[83-87] depression,[85, 88, 89] anxiety,[77, 90] negative affect,[80, 86] and expectations,[86] have all been implicated in the development and progression of chronic LBP. Other studies contradict these findings and suggest there is no link between psychological factors (e.g., fear-avoidance beliefs) and poor prognosis.[91] Distinguishing which psychological factors are most important in predicting the development of chronic LBP is not clear.[9, 57, 92-95]

One of the classic systematic reviews on the influence of psychological factors associated with spinal pain is provided by Linton (2000).[61] A comparison across settings and time points in 37 prospective studies indicated that stress, distress, anxiety, and

depressed mood were consistently related to future disability across settings (i.e., general population, primary care, secondary care, workplace) and time points (i.e., onset of pain episode, acute or subacute, chronic). One study in particular was relevant to this dissertation based on clinical setting and reported clinical outcomes.[96] The findings of that study indicated that elevated depression was highly related to future disability in patients with chronic LBP.[96]

One of the most recent systematic reviews involving psychological factors as prognostic indicators for persistent pain and disability is provided by Nicholas and colleagues.[11] A comparison across 12 prior review studies indicated consistent relationships between depression, pain catastrophizing, pain intensity, and beliefs about pain with future clinical or occupational outcomes in patients with acute or subacute LBP. Furthermore, the authors highlighted that many commonly used psychological screening instruments (e.g., FABQ, TSK) may be better suited for patients with persistent pain, therefore suggest potential benefits of using single, composite instruments with a small number of items (e.g., STarT Back Tool) for predicting risk of future clinical outcomes in physical therapy settings.[11]

One of the most recent systematic reviews involving predictors of poor clinical outcomes at 1-year is provided by Chou and Shekelle.[97] A comparison across 20 prospective studies consisting of patients with LBP of less than 8-weeks in duration (i.e., acute or subacute) indicated that nonorganic signs, elevated maladaptive pain coping behaviors, elevated baseline LBP-related disability, the presence of psychiatric comorbidities, and low general health status were the strongest predictors of poor clinical outcomes at 1-year follow-up. Alternatively, low levels of fear-avoidance and low

baseline LBP-related disability were strongest predictors of recovery at 1-year follow-up. Work environment, baseline pain intensity, and presence of radiculopathy were not as useful for predicting poor clinical outcomes, while LBP episode history and demographic variables were not at all useful.[97] The authors suggested that because individual risk factors were relatively weak, risk prediction screening instruments could be more useful in predicting poor clinical outcomes in comparison to screening instruments focused on a single domain.

In summary, previous review studies have provided evidence that elevated measures of several different psychological factors are positively associated with poor future clinical outcomes in a variety of clinical settings, including physical therapy, however there is a definite need for improving the methodology of primary and review studies involving the prognosis of LBP.[57] Relevant to this dissertation, potential benefits of incorporating brief, composite risk prediction instruments for future clinical outcomes (e.g., STarT Back Tool) have been suggested.[11, 13, 98]

Measuring the Influence of Psychological Factors with Self-Report Questionnaires

Screening Process

The intent of primary prevention is the protection of health by personal and community-wide efforts. As a potential component of primary prevention, screening can provide valuable information regarding risk factors for future disease among healthy individuals in the general population (e.g., demographics or lifestyle).[99] However, screening is more commonly associated with secondary prevention processes where the intent is early identification of individuals with the potential for poor future outcomes (e.g., LBP-related disability).[98] As previously described, “psychologically informed

practice" has been presented as a secondary prevention approach to physical therapy for chronic LBP that integrates both biomedical (focused on physical impairments) and cognitive-behavioral (focused on psychological distress) intervention principles.[13] The "flag" system has been suggested as a framework to classify patients and assist in clinical decision-making processes based on flag colors representing different types of risk factors.[11, 13]

- Red Flags – serious pathology (e.g., fracture)
- Orange Flags – psychopathology (e.g., clinical depression)
- Yellow Flags – normal psychological reactions to symptoms (e.g., fear-avoidance beliefs about physical activity)
- Blue Flags – perceptions about work and health relationships (e.g., belief that increased work will lead to further injury)
- Black Flags – healthcare system influence on clinical decisions and contextual factors (e.g., insurance restrictions, socioeconomic status)

At this point, it is also important to distinguish between modifiable and non-modifiable psychological risk factors from a physical therapy intervention perspective because both may be strong predictors of poor future outcomes identified through screening. Main and George[13] suggest that the ability to distinguish between these two types of risk factors based on the flag system is a critical component to psychologically informed practice because physical therapists are not trained to address all psychological risk factors. For example, if properly trained, physical therapists are equipped to provide interventions tailored to addressing yellow flags (e.g., fear-avoidance beliefs); which are considered modifiable psychological risk factors. However, it is not within the scope of physical therapy practice to provide intervention for orange flags (e.g., clinical depression); which are considered non-modifiable psychological risk factors through physical therapy intervention – however can be treated by other healthcare professionals (e.g., clinical psychologists).

Previous suggestions indicate that implementing appropriate treatment (e.g., behavioral interventions) during early stages of the LBP experience may be advantageous because of the adverse effects that psychological factors may exert over a prolonged period.[8, 100, 101] Results of a high-quality RCT also suggest that cognitive-behavioral approaches may be detrimental for individuals not deemed appropriate for this treatment approach.[102] There is an adequate amount of literature to suggest that psychological risk factors are modifiable with appropriate interventions and more importantly that psychological risk factor modification is associated with improved patient clinical outcomes.[102-108] Therefore, the rationale for screening of modifiable psychological risk factors is clinically important because of the potential implications these factors may have on establishing a prognosis and in implementing appropriate interventions. Although the processes involved with establishing a prognosis and screening share similarities, they should be interpreted as separate entities in the assessment of patients with LBP. Screening procedures can be used in the early identification of factors that may be influential in the development of chronic disability.[98] Additionally, the purpose of psychological screening in patients experiencing LBP has been suggested to be useful in matching patients with appropriate interventions prior to the development of chronicity.[100, 109] In comparison, the process involved with establishing a prognosis is often more comprehensive in nature and consists of findings from the patient's history and diagnostic tests during the physical examination. Therefore, screening tests should not be interpreted as diagnostic tools,[110] however can be considered "a rough assessment to narrow down the number of patients who need to be assessed in more detail".[100]

The literature involving screening for modifiable psychological risk factors is extensive and includes studies that use self-report questionnaires ranging from those focusing on a single-construct that measure the influence of a specific psychological construct to those that utilize a multiple-construct approach which determine risk based on overall psychological distress. There are strengths and weaknesses associated with self-report questionnaires that incorporate either single or multiple psychological construct approaches. For example, single-construct screening questionnaires do not investigate the potential influence that other psychological constructs may have on the risk of future clinical outcomes unless several questionnaires are used, which is often not feasible in clinical settings. Multiple-construct questionnaires may be more feasible to use in clinical settings, however are frequently used to determine risk based on overall psychological distress and do not provide detailed information on specific psychological constructs that may ultimately be treatment targets. This is an important issue because information on specific psychological constructs may be needed to implement psychologically oriented interventions that reduce the likelihood of developing chronic LBP. Therefore, an optimal clinical scenario may consist of screening with a brief multiple-construct risk prediction instrument (e.g., STarT Back Tool) to determine which patients require further detailed measurement via individual questionnaires.[11, 13, 98]

Studies of Psychological Measures: Single Construct Approaches

Single-construct measurement approaches can be used to quantify the influence of a single, specific psychological construct of interest via the use of questionnaires containing a single item or multiple items. The following study results are focused on

comparing the contributions of specific psychological factors when using a single-construct measurement approach.

Crombez and colleagues[74] reported on the results of 3 independent studies investigating different aims. In these studies, pain-related measures were correlated with each other (r = 0.34 to 0.76) and pain catastrophizing (r = 0.53 to 0.61; study 3), which may suggest construct redundancy of total scores. Specifically, pain-related fear measures were better predictors of self-reported disability (FABQ-PA, FABQ-W, TSK (17-item version); standardized β range = 0.40 to 0.57) and physical impairment (FABQ-PA, TSK; standardized β range = -0.27 to -0.28) than pain intensity ratings or negative affect in patients with chronic LBP.[74] Results from a separate analysis in the same cohort of patients suggested that pain-related fear as measured by the TSK was a better predictor of disability and physical impairment in comparison to negative affect or pain catastrophizing.[74]

McCracken and colleagues[76] compared several FAM measures (PASS, STAI-trait, FABQ-PA, FABQ-W, and FPQ) in their ability to predict baseline clinical characteristics in a cohort of chronic pain patients (LBP – 62%) referred to a pain clinic. Although this study consisted of a limited sample size (n = 45), results suggested that anxiety as measured by the PASS, accounted for 16% to 54% of the variance in pain severity, perceived disability (with FABQ-W), and pain behavior models (with FABQ-W).[76]

Fritz and colleagues[84, 111] reported on the importance of fear-avoidance beliefs as measured by the FABQ in patients with acute work-related LBP. Results suggested that fear-avoidance beliefs as measured by the FABQ-W was a significant predictor of disability and work status following 4-weeks of physical therapy, while depression as

measured by the CES-D and anxiety as measured by the Beck Anxiety Index were not.[84, 111] Similar results have been reported in other studies involving work-related LBP.[112]

de Souza and colleagues[113] reported that TSK (17-item version) and FABQ scores were highly correlated with each other (r = 0.86) when Portuguese versions of these measures were administered in a cohort of LBP patients. Both measures were also moderately correlated with pain intensity (r = 0.42 to 0.43). TSK scores were moderately correlated with Global Perceived Effect (GPE) scale scores (r = -0.46), while FABQ scores were not. Moreover, the TSK outperformed FABQ total and subscale scores in identifying change in GPE scores overtime.[113]

Woby and colleagues[114] reported that several psychological measures (FABQ-PA, FABQ-W, catastrophizing (via CSQ subscale), appraisals of control (via CSQ – 2 single-items)) were significantly correlated with each other (r = -0.37 to 0.51), with the exception of FABQ-W and appraisals of control over pain (r = -0.10, $p > 0.05$) in a cohort of chronic LBP patients. Hierarchical regression analyses revealed that only a greater ability to decrease pain (appraisal of control) ($\beta = -0.24$, $p < 0.05$) contributed a small but statistically significant proportion of the variance (6%) in pain intensity. After adjusting for age, sex, and pain intensity, only FABQ-PA ($\beta = 0.39$, $p < 0.01$) made a unique contribution to the prediction of disability in a final model that explained 52% of the variance in disability.[114]

Meyer and colleagues[73] reported that several psychological measure scores (PCS, FABQ-PA, FABQ-W, Modified Somatic Perception Questionnaire, Modified Zung Depression Scale) were significantly correlated with pain intensity ratings (r = 0.23 to

0.60), disability scores (r = 0.52 to 0.70), and each other (r = 0.28 to 0.61) in a cohort of chronic LBP patients. In the pain model, only FABQ-W (β = 0.40, p < 0.01) and modified somatic perception (β = 0.35, p < 0.01) made unique contributions in a model that explained 42% of variance in pain intensity ratings. In the disability model, only FABQ-W (β = 0.40, p < 0.01) and depression (β = 0.24, p < 0.05) made unique contributions in a model that explained 59% of variance in disability scores.[73]

George and colleagues[78] compared several FAM measures (FABQ-PA, FABQ-W, FPQ-9, TSK-11, PCS) in a physical therapy setting. Similar to previous studies, FAM measures were significantly correlated with each other (r = 0.30 to 0.69), with the exception of FPQ-9 and FABQ-W (r = 0.04, p>0.05). Criterion validity of these FAM measures was examined by their ability to predict baseline clinical outcomes. After controlling for age, sex, and employment status in separate multiple regression models: 1) the PCS contributed an additional 37% (p <0.01) variance in depression scores; 2) the FABQ-PA contributed an additional 23% (p<0.01) variance in pain intensity ratings; 3) the FABQ-W contributed an additional 13% (p<0.01) variance in physical impairment scores; and 4) the FABQ-PA and FABQ-W contributed an additional 23% and 8% variance in disability scores.[78]

Lundberg and colleagues[115] investigated the individual contribution of FAM variables (i.e., pain intensity, kinesiophobia, depressed mood) on baseline disability (i.e., ODQ scores) for patients with specific or nonspecific chronic LBP in an orthopaedic outpatient setting using a cross-sectional study design. For patients with specific LBP, after controlling for age and sex, FAM variables explained 67% of the variance in disability scores with each contributing uniquely (pain intensity (β = 0.48),

kinesiophobia (β = 0.18), depressed mood (β = 0.42)) For patients with nonspecific LBP, after controlling for age and sex, FAM variables explained 63% of the variance in disability scores with only pain intensity (β = 0.51) and depressed mood (β =0.40) contributing uniquely.[115]

Foster and colleagues[116] investigated the ability of 20 psychological factors to predict 6-month disability (i.e., RMDQ scores) following primary care consultation using a longitudinal study design. Five previously validated screening instruments were used in this study (i.e., the Revised Illness Perception Questionnaire; the Tampa Scale of Kinesiophobia; the Coping Strategies Questionnaire; the Hospital Anxiety and Depression Scale; the Pain Self-Efficacy Questionnaire). Only 2 of the baseline psychological factor scores were not significantly associated with baseline disability scores (i.e., illness perceptions about timeline – cyclical, coping subscale – interpretation). The 20 factors accounted for between 0.04% and 33.3% of the variance in baseline disability scores. A univariate analysis resulted in 11 factors that were associated with disability scores at 6-months (i.e., perceptions about consequences, emotional representations, personal control, treatment control, timeline – acute/chronic, illness identity, immunity attribution, depression, pain self-efficacy, fear-avoidance, and catastrophizing). However, after controlling for demographic and baseline clinical characteristics, psychological factors only accounted for 0.5 to 4.9% ($p < .01$) additional variance in disability scores at 6-months.[116] Final multivariate analyses resulted in only 4 factors (i.e., perceptions of personal control (β = -0.10), timeline – acute/chronic (β = 0.20), illness identity (β = 0.07), and pain self-efficacy (β = -0.11)) that remained statistically significant ($p < .01$) in a final model explaining 56.6% of LBP-related

disability. In this model, depression, fear-avoidance, and catastrophizing were no longer statistically significant.[116] The authors stressed the importance of illness perceptions and self-efficacy as psychological factors which may predictive of future outcome. Based on the results of this study, others have commented that other "usual suspects" (e.g., depression, fear-avoidance, and catastrophizing) "are not always guilty".[117]

In summary, many FAM-focused psychological factors are positively correlated with each other when comparing scores obtained from single-construct screening questionnaires. Collectively, FAM-focused psychological factors are strongly linked to baseline clinical characteristics and future clinical outcomes; however the magnitude of these relationships is not always consistent.

Studies of Psychological Measures: Multiple Construct Approaches

Multiple-construct measurement approaches can be used to quantify general psychological distress via the use of risk prediction instruments by combining responses from several single-item questions, each representing a different psychological construct. The following study results are focused on the STarT Back Tool for the purpose of this dissertation; however it is acknowledged that similar risk prediction instruments have been reported in the literature (e.g., the Vermont Disability Prediction and Örebro Musculoskeletal Pain Screening Questionnaires).[118, 119]

The STarT Back Screening Tool which is a 9-item screening questionnaire that consists of items relating to referred leg pain, comorbid pain, disability, and psychological factors (i.e., bothersomeness, catastrophizing, fear, anxiety, and depression).[120] Based on STarT responses, patients are allocated into 1 of 3 subgroups (i.e., low, medium, and high risk) describing risk status for future disability

and are associated with initial treatment options in primary care settings. For example, patients allocated to the medium risk subgroup are deemed suitable for standard physical therapy intervention, whereas those allocated to the high risk subgroup may require a combination of physical and cognitive-behavioral treatment approaches.[120] The STarT was developed and validated in the primary care setting and was intended to be utilized for the identification of treatment subgroups in that setting.

Specific to physical therapy settings, a previous prospective study conducted by our group involving patients (n = 214) seeking care for LBP in the outpatient physical therapy setting (n = 3) reported on the distribution of STarT subgroups (33% low risk; 48% medium risk; 19% high risk).[121] Patients allocated to STarT low risk had lower initial pain and disability scores in comparison to those allocated to STarT medium or high risk ($p < .001$) and patients allocated to STarT high risk had higher initial pain and disability scores in comparison to patients allocated to STarT medium or low risk ($p < .001$). In addition, patients allocated to STarT medium risk had higher initial pain and disability scores in comparison to patients allocated to STarT low risk ($p < .001$). Further analyses using linear mixed modeling techniques were used to investigate patterns of change in predicted clinical outcomes across the episode of care in a subset of patients (n = 177). Relative to STarT low risk, the high risk subgroup had larger improvements in predicted outcomes and the medium risk subgroup had similar improvements. Variable timing of follow-up assessments was a potential limitation to this study.[121] Additional studies that have reported on STarT Back Tool psychometric properties are presented in the methods section of this dissertation.

Public Health Significance and Relationship to Rehabilitation Science

Previous epidemiological literature has indicated that LBP is a major public health problem at both the individual and societal levels with the development and progression to chronic LBP aggravating matters and accounting for a majority of healthcare resources. Previous healthcare service literature has indicated that physical therapy services compromise one of the largest proportions of direct medical costs for LBP,[6] which is not surprising considering nearly one-third of all patients in outpatient physical therapy settings seek care for LBP.[26] Specific to this dissertation, the FAM provides an explanation for why a minority of individuals that experience an acute episode of LBP eventually develop chronic LBP that is primarily influenced by psychological factors. Psychological screening is primarily focused on modifiable risk factors. This is relevant for physical therapists because early identification of modifiable psychological risk factors for poor clinical outcomes may be most appropriate for patients that have not yet developed and progressed to chronic pain states (i.e., secondary prevention).[11, 13] Furthermore, it is important that clinicians in outpatient physical therapy settings are able to distinguish between “yellow flags” (e.g., catastrophizing, fear-avoidance beliefs) and “orange flags” (e.g., depression, anger) with the former appropriate for physical therapy intervention and the latter requiring consultation from mental health professionals.[11, 13] Focusing on routine and specific methods to identify the influence of psychological factors on clinical outcomes has been recommended by Main and George as a component of psychologically informed practice for the management of LBP in physical therapy settings.[13] Therefore, a viable goal for rehabilitation scientists is to improve the screening process for patients with LBP in order to detect those at risk for developing chronic symptoms. Determining optimal screening procedures will provide

clinicians with valuable prognostic and/or treatment effect modification information to assist in clinical decision-making algorithms that may lead to better clinical outcomes for patients with elevated psychological distress.

CHAPTER 3
RESEARCH HYPOTHESES

The primary goals of this dissertation were to: 1) determine the validity of the STarT Back Tool (a multi-construct psychological screening measure) in outpatient physical therapy settings and 2) test the clinical utility of the STarT Back Tool in comparison to commonly used single-construct psychological screening measures. These goals were addressed through three specific aims. The following sections provide initial hypotheses and support for each specific aim.

Specific Aim 1

Construct Validity of the STarT Back Tool Classification Scheme at Initial Physical Therapy Evaluation

We hypothesized that patients allocated to the STarT high risk subgroup would be associated with higher baseline pain intensity, disability, physical impairment, and psychological distress in comparison to patients allocated to STarT medium and low risk subgroups. A previous study by our group in an outpatient physical therapy setting supports this hypothesis for baseline pain intensity and disability scores, however did not test for relationships between STarT risk allocation and physical impairment or psychological scores.[121] A previous study conducted in a chiropractic setting supports this hypothesis for several individual FAM-based psychological measures (i.e., fear-avoidance beliefs, pain catastrophizing, and depressive symptoms.[122]

Discrimination of Psychological Factors amongst STarT Back Tool Classification Status at Initial Physical Therapy Evaluation

For this exploratory aim, we hypothesized that depressive symptoms (PHQ-9) would demonstrate a strong ability to discriminate initial STarT classification status, while measures focused on pain-related fear (PCS, TSK-11) and fear-avoidance beliefs

(FABQ-W, FABQ-PA) would demonstrate moderate abilities. These hypotheses were primarily based on collective results from prior studies using similar subgrouping methodologies. Separate studies using a screening instrument similar to the STarT Back Tool indicated that groups defined as “pain-related fear and depressed mood” and “distressed fear-avoidant” were associated with poor initial clinical characteristics and outcomes in comparison to “low psychological distress” groups.[100, 123] A previous unpublished study by our group (in revision) using discriminant function analyses, indicated that FABQ-W, PCS, and FDAQ scores demonstrated strong relationships, while FABQ-PA scores were associated with weaker relationships to three cluster solutions based on the FAM (i.e., low risk, high specific fear, and high fear and catastrophizing).

Clustering of Psychological Factors at Initial Physical Therapy Evaluation without considering STarT Back Tool Classification Status

For this exploratory aim, we hypothesized that two to four distinct FAM-based psychological profiles would result from a cluster analysis. Again, this hypothesis is primarily based on a previous unpublished study by our group (in revision) that used similar subgrouping methodologies to create three distinct cluster profiles (i.e., low risk, high specific fear, and high fear and catastrophizing) amongst four FAM measures (i.e., FABQ-PA, FABQ-W, FDAQ, and PCS). We further hypothesized that our emerging cluster profiles would be similar to the STarT risk categorization scheme based on psychological distress.

Specific Aim 2

Predictive Validity of the STarT Back Tool (4-Week Outcomes)

We hypothesized that STarT psychosocial scale scores measured on a continuous scale would have similar ability to predict 4-week clinical outcomes when compared to individual FAM based single-construct psychological measures in physical therapy settings. Support for this hypothesis is primarily based on the STarT Back Tool development study from a primary care setting where STarT psychosocial scale scores best discriminated (via area under the receiver operating characteristic curve (AUC) estimates) catastrophizing (0.83), fear (0.81), depressive symptoms (0.76), and disability (0.90) reference standards.[120] Furthermore, results from a separate study from primary care indicated that STarT psychosocial subscale scores were highly correlated with fear ($r = 0.66$) and catastrophizing ($r = 0.67$) scores.[124]

Predictive Validity of the STarT Back Tool (4-Week Outcome Change Scores)

For this exploratory aim, we hypothesized that results would be similar to those described above for 4-week outcomes, however wanted to investigate potential differences based on multivariate regression modeling techniques. Specifically, we intended to investigate for differences in results if: 1) initial scores were controlled for in our models with 4-week outcomes serving as the dependent variable or 2) 4-week outcome change scores served as dependent variables and initial scores were not entered into our models.

Specific Aim 3

Prediction of Sustained STarT High Risk Allocation

For this aim, we hypothesized that FABQ-PA, FABQ-W, and PCS scores would be best at predicting patients allocated to STarT high risk at initial physical therapy

evaluation and remained high risk at 4-weeks. Preliminary support for this hypothesis is primarily based on: 1) cross-sectional studies investigating relationships between individual FAM based psychological measures and clinical characteristics and 2) the STarT Back subgrouping scheme associating high-risk patients with poor future outcomes. Previous studies by our group [78, 79] provide preliminary support for this hypothesis by indicating that initial FABQ-PA, FABQ-W, and PCS scores were predictive of initial pain and disability scores in outpatient physical therapy settings.

Sustained STarT High Risk Allocation and Outcomes

For this aim, we hypothesized that patients allocated to STarT high risk at intake and remained high risk at 4-weeks would have poorer outcomes compared to patients allocated to STarT high risk at intake and changed to low or medium risk at 4-weeks.

Relevance and Novelty of Specific Aims

The relevance and novelty of this dissertation can be demonstrated in several ways. First, the STarT Back Tool was developed and intended for use in primary care settings.[120] To date there has only been one published study reporting on data from patients in outpatient physical therapy settings,[121] therefore the results of this study have potential to impact physical therapy practice, which is both a relevant and novel aspect of this dissertation. Second, in the attempt to meet previous research priorities, [5, 11, 13, 56] the STarT Back Tool (a multiple-construct risk prediction instrument) was compared to individual FAM based psychological measures using a longitudinal study design. Incorporating these previous research priorities into the design of this study is a relevant aspect of this dissertation that has potential to positively impact future physical therapy research and practice when standardized psychological screening procedures are identified for patients with LBP. Third, while appealing to clinicians, there may be

disadvantages in using dichotomous cutoff scores of psychological measures for predicting risk of poor outcomes,[125] therefore analysis of continuously distributed data has been recommended which can be considered an additional relevant aspect of this dissertation.[13] Fourth, this study design consisted of standardized timing of follow-up assessments and measured multiple clinical outcome domains. These are important criteria that have been recommended for future research[126] and are relevant based on difficulties when attempting to compare results of previous studies consisting of heterogeneous methodologies. Finally, modifiable psychological risk factors have been targeted in this dissertation with the next step being able to match patients with appropriate physical therapy interventions to address these factors which is both a relevant and novel aspect of this dissertation. Collectively, all of the above mentioned aspects of this dissertation are in line with a recommended shift in physical therapy management for LBP consistent with psychologically informed practice.[13]

CHAPTER 4
METHODS

Research Design

This study was designed to investigate the validity and clinical utility of the STarT Back Screening Tool in comparison to FAM based single-construct psychological measures. Consecutive patients seeking outpatient physical therapy services for LBP were considered for study participation (Figure 4-1). Patients meeting study inclusion criteria and who provided informed consent completed self-report forms for demographic, clinical, and psychological measures and underwent a standard physical examination that was performed by a licensed physical therapist. Then, patients received physical therapy intervention for LBP that was left to the discretion of the physical therapist. At 4-weeks following the initial evaluation, patients completed self-report forms for clinical and psychological measures and underwent a standard physical examination. In the event that patients were not able to attend the 4-week follow-up session in person, the option to complete self-report forms for clinical and psychological measures through mail was offered as an alternative. In these cases, a standard physical examination was not performed at 4-weeks.

Participants

Consecutive patients seeking outpatient physical therapy services for LBP in six outpatient physical therapy clinics located in Gainesville and Jacksonville, Florida were screened for study eligibility by a physical therapist.

Inclusion Criteria

Potential study participants met both of the following criteria before being enrolled into this study: 1) adults between the ages of 18 and 65 years seeking physical therapy for LBP and 2) the ability to read and speak the English language.

Exclusion Criteria

Potential study participants were ineligible to participate in this study if any of the following criteria were met: 1) the presence of systemic involvement related to metastatic or visceral disease; 2) recent fracture; 3) osteoporosis; or 3) pregnancy.

Procedures

Consecutive patients seeking outpatient physical therapy services for LBP in six outpatient physical therapy clinics located in Gainesville and Jacksonville, Florida were screened for study eligibility by a physical therapist at the initial evaluation session. Physical therapists provided all patients that met study eligibility criteria with a brief explanation of the study and a study advertisement with primary investigator contact information. Clinicians emphasized to patients that participating in this study would not dictate the treatment they received for their LBP and if they elected not to participate they would receive the same treatment. If appropriate, informed consent was obtained in compliance with the University of Florida's Internal Review Board. The following demographic, clinical and psychological measures were administered within the first two physical therapy sessions and again 4-weeks later.

Demographic and Clinical Measures

Study participants were asked to complete a standardized self-report questionnaire consisting of questions related to age, sex, race, ethnicity, education, household income, marital and employment status. Additionally, information involving

LBP clinical characteristics (i.e., prior surgery, symptom duration, symptom onset, symptom location, work-related LBP) was obtained.

Subgroups for Targeted Treatment (STarT) Back Screening Tool

The primary measure of interest for this dissertation is the STarT Back Tool which is a 9-item screening measure used to identify subgroups of patients with LBP in primary care settings based on the presence of potentially modifiable prognostic factors which may be useful in matching patients with targeted interventions.[120] The STarT contains items related to physical (items 2, 3, 5, and 6) and psychosocial (items 1, 4, 7, 8, and 9) factors that have been identified as strong independent predictors for persistent disabling LBP. Potential responses for the STarT are dichotomized ('agree' or 'disagree'), with the exception of an item related to 'bothersomeness' which uses a 5-point Likert scale. Overall STarT scores (ranging from 0 to 9) are determined by summing all positive responses. Psychosocial subscale scores (ranging from 0 to 5) are determined by summing items related to bothersomeness, fear, catastrophizing, anxiety, and depression (i.e., items 1, 4, 7, 8, and 9). Based on overall and psychosocial subscale scoring, the STarT categorizes patients as 'high-risk' (psychosocial subscale scores ≥4) in which high levels of psychosocial prognostic factors are present with or without physical factors present, 'medium-risk' (overall score >3; psychosocial subscale score <4) in which physical and psychosocial factors are present, but not a high level of psychosocial factors, or 'low-risk' (overall score 0-3) in which few prognostic factors are present.[127]

The STarT overall (0.79, 95% CI: 0.73 – 0.95) and psychosocial subscale (0.76, 95% CI: 0.52 – 0.89) scores has been found to have acceptable test-retest reliability (weighted kappa values) in patients with stable symptoms.[120] Cronbach's alpha

estimates for overall (0.79) and psychosocial subscale (0.74) scores suggest the STarT has demonstrates internal consistency.[120] The predictive validity of the STarT has been reported in which subgrouping cutoff scores were predictive of poor 6-month disability outcomes in low (16.7%), medium (53.2%), and high (78.4%) risk subgroups.[120] The discriminant validity of the STarT scores (AUC range: 0.73 – 0.92) have been reported and suggest that overall scores best discriminate physical reference standards (e.g., disability and referred leg pain), while psychosocial subscale scores best discriminate psychosocial reference standards (e.g., catastrophizing, fear, and depression).[120] The STarT has demonstrated concurrent validity in comparison to the Örebro Musculoskeletal Pain Screening Questionnaire, in which both instruments displayed similar subgroup characteristics and the ability to discriminate for disability, catastrophizing, fear, comorbid pain and time off work reference standards.[124] Danish and Spanish versions of the STarT Back Tool have been validated.[128, 129] Results from a cross-sectional study in the chiropractic setting (n = 475) reported on the distribution of STarT subgroups (59% low risk; 29% medium risk; 11% high risk) and dose-response relationship for STarT subgroup status with continuous scores for depressive symptoms (Major Depression Inventory), fear-avoidance beliefs (FABQ – total scores) and catastrophizing (CSQ – catastrophizing subscale).[122] A summary of STarT subgroup distributions across studies by setting is presented in Figure 4-2.

Fear-Avoidance Model Based Psychological Measures

Fear-Avoidance Beliefs Questionnaire

Fear-avoidance beliefs specific to LBP were assessed with the FABQ.[130] The FABQ consists of a 4-item FABQ physical activity scale (FABQ-PA, potentially ranging from 0 to 24) and a 7-item FABQ work scale (FABQ-W, potentially ranging from 0 to

42), with higher scores indicating higher levels of fear-avoidance beliefs for both FABQ scales. Patients rated their agreement with statements related to either physical activity or work on a 7-point Likert scale (0 = "completely disagree," 6 = "completely agree").[130]

The FABQ scales have been found to have acceptable reliability.[130-133] Test-retest reliability has been reported for the FABQ-PA (Pearson r = 0.84 to 0.88) and FABQ-W (Pearson r = 0.91 to 0.88).[130, 133] Cronbach's alpha estimates for the FABQ-PA (ranging from 0.70 to 0.83) and FABQ-W (ranging from 0.71 to 0.88) scores suggest both scales demonstrate internal consistency.[81, 130, 133-135] The FABW has demonstrated predictive validity for disability and work loss in patients with LBP.[84, 102, 111, 133] A suggested FABQ-W cutoff score of >29 has been suggested as an indicator of return to work status in patients receiving physical therapy for acute occupational LBP[111] and a cutoff score of >22 has been suggested in non-working populations.[136] An FABQ-PA cutoff score of >14, based on a median-split of the FABQ has been suggested as an indicator of treatment outcomes in LBP patients seeking care from primary care or osteopathic physicians.[105] George and colleagues analyzed data from 2 separate physical therapy intervention clinical trials and found that the FABQ-W cutoff score (>29) was a better predictor of self-reported disability at 6-months in comparison to the FABQ-PA cutoff score (>14).[136]

Pain Catastrophizing Scale

The PCS was used to assess the degree of catastrophic cognitions due to LBP.[137] Pain catastrophizing has been broadly defined as an exaggerated negative orientation towards actual or anticipated pain experiences.[137] The PCS is a 13-item questionnaire with a potential range of 0 to 52, with higher scores indicating higher levels of pain catastrophizing. Patients rated their agreement with statements related to thoughts and

feelings when experiencing pain on a 5-point Likert scale (0 = “not at all,” 4 = “all the time”).[137] The PCS assesses 3 independent dimensions of pain catastrophizing: rumination (items 8-11 – ruminating thoughts, worrying, inability to inhibit pain related thoughts); magnification (items 6,7,13 – magnification of the unpleasantness of pain situations and expectancies for negative outcomes); and helplessness (items 1-5, 12 – inability to deal with painful situations).[137, 138]

Test-retest reliability has been reported for the PCS at 6 (r = .75) and 10-weeks (r = .70).[137] Cronbach’s alpha estimates ranging from .85 to .92 suggest the PCS is internally consistent.[74, 139, 140] and similar findings have been found for items related to rumination (.85), magnification (.75), and helplessness (.86).[140] The PCS has been found to demonstrate several different types of validity.[74, 137, 139, 140]

Tampa Scale of Kinesiophobia

The TSK-11 was used to assess the degree of fear of movement and injury or re-injury in individuals with LBP.[141] The TSK-11 is an 11-item questionnaire with a potential range of 11 to 44, with higher scores indicating greater fear of movement and injury or re-injury due to pain. Patients rated their agreement with statements related to fear of movement and injury or re-injury when experiencing pain on a 4-point Likert scale (1 = “strongly disagree,” 4 = “strongly agree”).

Test-retest reliability (ICC = 0.81; 95% CI: 0.71 – 0.88) and internal consistency (Cronbach’s alpha = .79) have bee reported for the TSK-11.[141] The concurrent validity of this instrument has been reported in which changes in disability were correlated with TSK-11 change scores (r = 0.51).[141] The predictive validity of the TSK-11 has also been reported in which reductions on TSK-11 scores explained an additional 12% of variance in disability in patients with chronic LBP.[141] There have been no reports of

suggested cutoff scores to identify elevated levels of pain-related fear using the TSK-11, however a 4-point reduction in TSK-11 scores has been suggested as an indicator in identifying important reductions in fear of movement (sensitivity = 66%, specificity = 67%).[141]

Patient Health Questionnaire

The PHQ-9 was used to assess the degree to which depressive symptoms have on patients with LBP. The PHQ-9 is a 9-item questionnaire with a potential range of 0 to 27, with higher scores indicating elevated depressive symptoms. Patients rated their agreement with statements related to signs and symptoms of depression on a 4-point Likert scale (0 = "not at all," 3 = "nearly every day"). The PHQ-9 total score was used for this study. The PHQ-9 has been validated in different settings[142, 143] and has been used in studies involving patients with LBP.[144] In addition to being used as a screening instrument, the PHQ-9 has also been suggested for use as an outcome measure.[145] A PHQ-9 cutoff score of ≥10 has been suggested as an indicator of major depression (sensitivity = 88%, specificity = 88%; positive LR = 7.1).[142]

State-Trait Anxiety Inventory

The trait portion of the STAI (STAI-T) was used to assess the degree that dispositional anxiety has on patients with LBP.[146] Trait anxiety has been suggested to be more closely related to disability following episodes of musculoskeletal pain[90] and the STAI-T has been used in other studies involving LBP patients in physical therapy settings.[147] The STAI-T is a 20-item questionnaire with a potential range of 20 to 80, with higher scores indicating elevated levels of anxiety. Patients rated their agreement with statements related to signs and symptoms of trait anxiety on a 4-point Likert scale (1 = "almost never," 4 = "almost always"). The STAI-T has been found to be reliable and

valid,[148, 149] however the use of cutoff scores have resulted in less than optimal results in a sample of psychiatric outpatients.[149] A potential limitation in using the STAI-T is that it has been found to be limited in its ability to differentiate anxiety from depressive disorders.[150] Therefore, we used a validated depression screening instrument (i.e., PHQ-9) in this study.

Outcome Measures

Pain Intensity

Pain intensity was rated using a numerical pain rating scale (NPRS), ranging from "0" (no pain) to "10" (worst pain imaginable).[151-153] Participants were asked to rate their current pain intensity, as well as their best and worst level of pain intensity over the past 24 hours.

Revised Oswestry Disability Questionnaire

LBP-related disability was assessed with the revised Oswestry Disability Questionnaire (ODQ), which has 10 items that assesses how LBP affects common daily activities.[154, 155] The ODQ has a range of 0% "no disability due to LBP" to 100% "completely disabled due to LBP", with higher scores indicating higher disability from LBP. The ODQ has been found to have high levels of test-retest reliability, internal consistency, validity, and responsiveness.[155-157]

Roland-Morris Disability Questionnaire

LBP-related disability was also assessed with the Roland-Morris Disability Questionnaire (RMDQ), which has 24 items that assesses the functional status over the past 24 hours in patients with LBP.[158] The RMDQ has a range of 0 (no disability due to LBP) to 24 (maximum disability due to LBP), with higher scores indicating higher

disability from LBP. The RMDQ has been found to have high levels of test-retest reliability, internal consistency, validity, and responsiveness.[157, 158]

Physical Impairment Index

Physical impairment was assessed using the Physical Impairment Index (PII) which was used to establish an objective measurement of physical impairment in patients with LBP.[159] The PII consists of seven physical examination tests routinely implemented in a physical therapy physical examination for patients with LBP. The individual tests included: 1) lumbar flexion range of motion (ROM); 2) lumbar extension ROM; 3) lumbar lateral flexion ROM; 4) passive straight leg raise ROM; 5) bilateral active straight leg raise; 6) active sit-up; and 7) assessment of spinal tenderness. Each test was scored positive or negative based on published cut-off values.[159] The overall PII score ranges between "0" and "7", with higher scores indicating greater levels of physical impairment. Good or excellent reliability has been reported for individual items of the PII and convergent validity has been supported via correlations with disability in patients with chronic LBP.[159] Similar results have been reported in patients with acute LBP.[160] In patients with acute LBP, the overall PII score is more responsive to change than its individual test components, with a 1-point change representing a minimal detectable change over four-weeks of physical therapy.[160]

Sample Size Estimate

Since effect sizes for STarT tool from physical therapy settings were not available when this project was planned our sample size estimate (n = 150) was based on suggested guidelines (i.e., 10 cases per predictor variable) for creating multiple regression models that were not overfit.[161, 162] In our separate regression models for clinical outcomes (specific aim 2) there was the potential for 14 to15 variables to be

entered into each model. Therefore a liberal sample size estimate (n = 150) allowed us to meet these minimum guidelines for subsequent analyses.

Statistical Analyses

All data analyses were performed using SPSS, Version 18.0. Means and standard deviation were calculated for all baseline continuous variables and frequency counts with percentages were calculated for categorical variables. These descriptive statistics are presented for the entire study sample and for each STarT subgroup. The distributions of baseline continuous variables were examined by visual inspection of histograms and calculating skewness and kurtosis statistics. Nonparametric estimates were used for variables with significant deviations from a normal distribution based on visual inspection, skewness and/or kurtosis statistics.

Specific Aim 1

Construct validity of the STarT Back Tool classification scheme at initial physical therapy evaluation

The construct validity of the STarT was assessed by investigating how key initial demographical, clinical, and psychological variables discriminated amongst STarT subgroup status for patients seeking outpatient physical therapy. Demographic variables consisted of age, sex, race, ethnicity, employment status, education, and household income. Clinical variables consisted of LBP-associated: surgical status, symptom duration, symptom onset, symptom location, baseline pain intensity (NPRS average ratings), LBP-related disability (ODQ and RMDQ scores), and physical impairment (PII scores). Psychological variables consisted of measures of fear-avoidance beliefs (FABQ – physical activity and work scales), pain catastrophizing (PCS), pain-related fear (TSK-11), depression (PHQ-9), and anxiety (STAI-trait).

Baseline STarT risk subgroup differences in variables were tested with one-way analysis of variance (ANOVA) and post-hoc testing as appropriate for continuous variables and chi-square testing for categorical variables.

Discrimination of psychological factors amongst STarT Back Tool classification status at initial physical therapy evaluation

As an exploratory aim, we investigated how psychological factors discriminated amongst initial STarT Back subgroup status with a discriminant function analysis (DFA). DFA is a multivariate statistical procedure used to determine if a set of variables can predict group membership; in this case – initial STarT subgroup status. This procedure has been used in other studies investigating classification schemes for patients receiving physical therapy for LBP.[163] Specifically, we used discriminant function analysis to determine: 1) which psychological measures at intake differentiated STarT subgroup status at intake and 2) the accuracy in STarT subgroup allocation at intake using psychological measure scores at intake. Psychological variables (FABQ-PA, FABQ-W, PCS, TSK-11, STAI-T, PHQ-9) were entered into a DFA based on their contributions to: 1) the Fear-Avoidance Model and 2) STarT psychosocial subscale items that are focused on fear, catastrophizing, anxiety, and depression. This analysis included three groups (i.e., low, medium, and high-risk), therefore two discriminant functions were generated from the DFA. Eigenvalues were reported as a measure of variance, indicating how well the discriminant function discriminated between STarT subgroups with higher eigenvalues indicating greater discrimination. Canonical correlations were reported as a measure of the relationship between initial STarT subgroup status and the discriminant function, with chi-square tests used to determine

the significance of the relationship. Finally, a summary of classification results from the DFA was generated.

Clustering of psychological factors at initial physical therapy evaluation without considering STarT Back Tool classification status

As an additional exploratory aim, we investigated how psychological factors clustered at intake (without consideration for STarT subgroup status) with a cluster analysis. Raw scores for each psychological measure (i.e., FABQ-PA, FABQ-W, PCS, TSK-11, PHQ-9, and STAI-T) were transformed to z-scores to provide standardized scores for subsequent cluster analysis techniques. An exploratory hierarchical agglomerative cluster analysis was performed using Ward's clustering method with squared Euclidean distances as the similarity measure to create distinct cluster profiles among FAM measures. Agglomeration coefficients were inspected and plotted to establish the most optimal cluster solution based on the percent change between adjacent cluster solutions[164] and plot characteristics (i.e., elbow criterion).[165] To identify potential cluster group differences in demographic variables and baseline clinical measures, one-way ANOVA with Bonferroni post-hoc correction was used for continuous variables and chi-square analysis was used for categorical data.

Specific Aim 2

Predictive validity of the STarT Back Tool (4-week outcomes)

The predictive validity of the STarT Back Tool was compared to commonly used single-construct psychological measures for 4-week clinical outcomes. The ability of the STarT psychosocial subscale score as a prognostic factor was investigated with four separate multivariate hierarchical regression models with 4-week clinical outcomes (pain intensity (NPRS), LBP-related disability (ODQ, RMDQ), or physical impairment

(PII) scores) serving as dependent variables. These hierarchical regression models consisted of separate simultaneous blocks to account for baseline dependent variable scores (block 1 – baseline NPRS, ODQ, RMDQ, or PII scores depending on the model), demographic and clinical variables (block 2 – age, sex, household income, surgery for current condition, duration of current symptoms, and number of PT visits at 4-weeks), STarT scores (block 3 – baseline STarT psychosocial scores), and other psychological measure scores (block 4 – baseline FABQ-PA, FABQ-W, PCS, TSK-11, STAI-T, and PHQ-9 scores). Regression diagnostics were performed to assess for multicollinearity between predictor variables in all multivariate regression analyses.

Predictive validity of the STarT Back Tool (4-week outcome change scores)

As an exploratory aim, we tested the predictive validity of the STarT Back Tool in comparison to commonly used single-construct psychological screening measures for 4-week clinical outcome raw change scores. Our rationale for this separate analysis was that we wanted to investigate for potential differences in results based on different multivariate regression modeling techniques and because assessing change is a debatable issue when using self-report measures.[125] Four separate multivariate hierarchical regression models were constructed similar to those described for 4-week outcomes, with the following exceptions:1) raw change scores were calculated for each respective clinical outcome and served as the dependent variable by subtracting 4-week scores from initial scores, and 2) block1 – which consisted of baseline scores, was removed. Therefore, for models for this aim consisted of: block 2 – age, sex, household income, surgery for current condition, duration of current symptoms, and number of PT visits at 4-weeks; block 3 – baseline STarT psychosocial scores; and block 4 – baseline FABQ-PA, FABQ-W, PCS, TSK-11, STAI-T, and PHQ-9 scores.

Specific Aim 3

Prediction of sustained STarT high risk allocation

We investigated factors that were best at predicting patients allocated to STarT high risk at initial physical therapy evaluation and remained high risk at 4-weeks. The relationship between STarT subgroup status at 4-weeks and baseline demographic, clinical, and psychological variables was tested with hierarchical logistic regression analyses. STarT subgroup status at 4-weeks served as the dependent variable and was dichotomized as low / medium risk or high risk. We explored univariate relationships between potentially important baseline variables in our sample with STarT subgroup status at 4-weeks with zero-order correlations. The hierarchical logistic regression models consisted of separate simultaneous blocks to account for demographic and clinical variables (block 1) and baseline psychological measure scores (block 2) based on univariate relationships ($p < .15$); which is acknowledged as a liberal model-building approach.

Sustained STarT high risk allocation and outcomes

As an exploratory analysis, we investigated if patients allocated to STarT high risk at initial evaluation and remained high risk at 4-weeks demonstrated poorer outcomes compared to patients allocated to STarT high risk at intake and changed to low or medium risk at 4-weeks. We tested for group x time statistical interactions for all clinical outcomes using 2x2 repeated measures model analysis of variance (ANOVA) with between factors consisting of STarT status at 4-weeks (i.e., low/medium or high risk) and time factors consisting of intake and 4-week clinical outcome scores. These exploratory analyses only consisted of patients allocated to STarT high risk at intake.

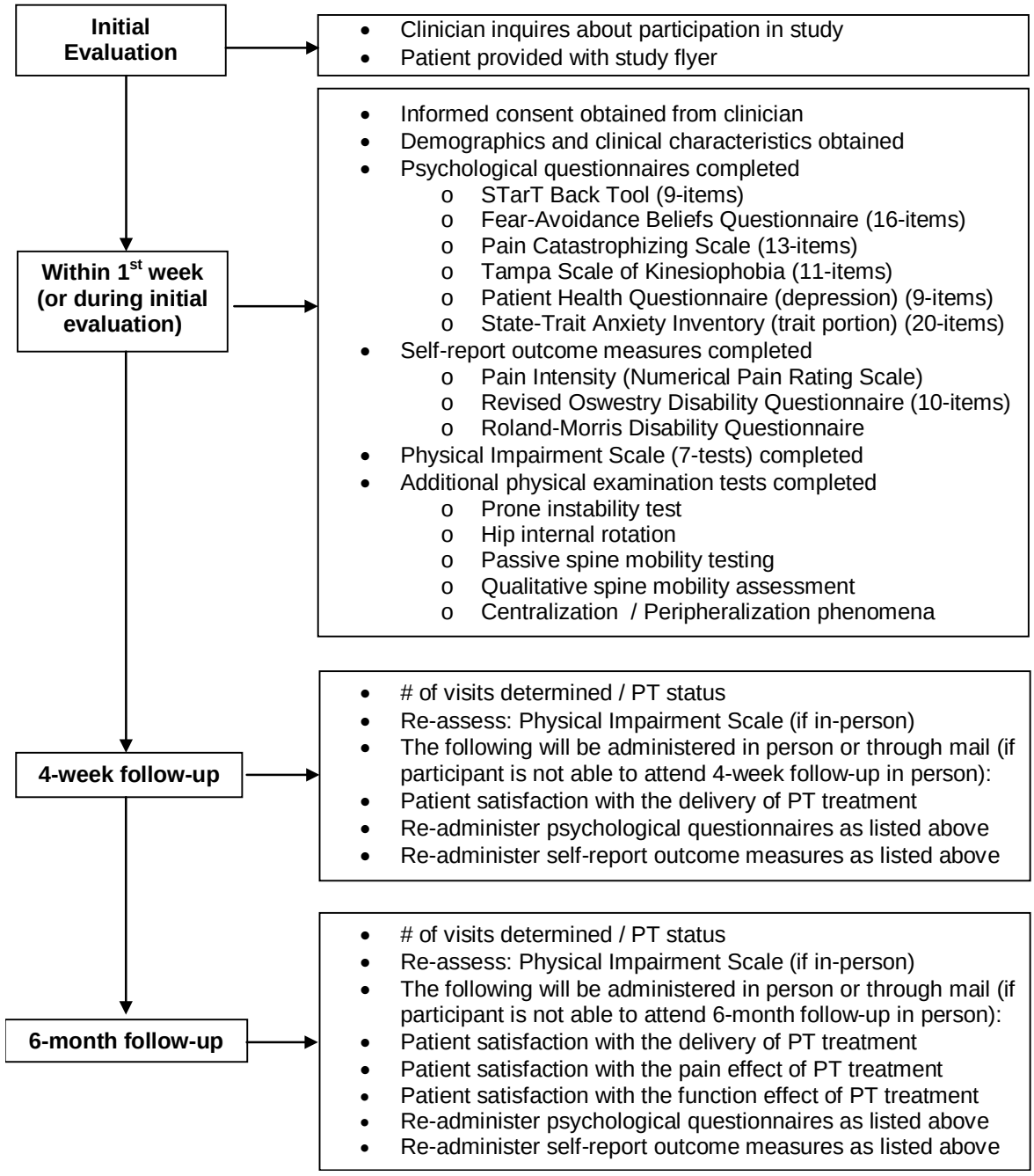

Figure 4-1. Flow chart of study design.

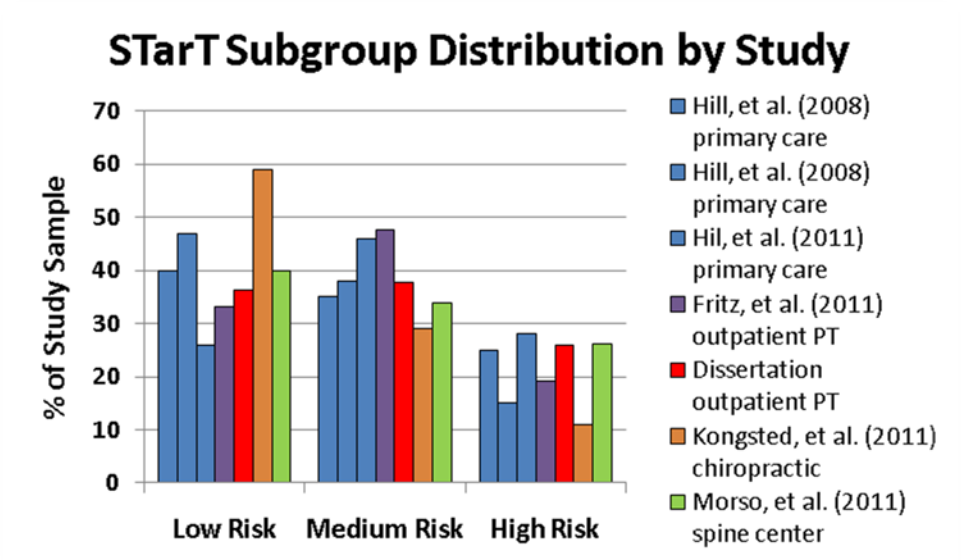

Figure 4-2. Initial STarT subgroup distribution across studies by setting.

CHAPTER 5
RESULTS

Participants

Between December 2009 and August 2011, 16 physical therapists from six different outpatient clinics recruited consecutive patients seeking physical therapy services for LBP. During this time 275 patients were screened for eligibility criteria. Of these patients, 123 were excluded from study participation with the most common reason being that they were greater than 65 years of age (n = 47). The remaining 152 patients provided informed consent and were enrolled into the study. Of these patients, six were dropped from the study due to personal reasons. Therefore, baseline data was obtained from 146 patients and 4-week follow-up data was obtained from 127 patients (87%) either in-person (n = 108) or through mail (n = 19).

Independent samples t-tests were used to compare patients that completed the 4-week follow-up assessment (n = 127) to those that did not (n = 19) on demographic, clinical, and psychological characteristics at initial evaluation. Results indicated that completers were older (mean difference = 7.9 years; SE = 3.2) and reported higher PHQ-9 scores (mean difference = 3.0; SE = 1.5) compared to non-completers (p's < .05). Furthermore, non-completers had a higher proportion of African-Americans (28.0%) compared to Caucasians (10.0%) or other races (9.1%) (p = .05).

Baseline demographic data for entire sample is presented in Table 5-1. The mean age of study participants was 41.4 (sd = 13.5) years of age, 89 (61.0%) were women, 26 (17.8%) had prior LBP-related surgery, and 19 (13.0%) experienced work-related LBP. The mean duration of LBP symptoms for the current episode of LBP was 483.5 (sd = 1157.7) days, with 11.6% acute (≤ 14 days), 38.4% subacute (15 to 90 days), and

47.9% chronic (> 91 days) episodes. Approximate normal distribution for initial pain intensity, disability, physical impairment, and individual FAM based psychological measures was suggested based on visual inspection of histograms and normal q-q plots.

The average number of physical therapy visits at 4-weeks was 6.87 (sd = 2.65; range = 1 to 13) where 100 (68.5%) participants were still receiving physical therapy, 15 (10.3%) completed physical therapy and were discharged, and 10 (6.8%) elected not to continue with physical therapy. There were no differences in initial or 4-week pain intensity, disability, physical impairment, psychological measure scores, or STarT subgroup status between participants that participated in 4-week follow-up assessments in person compared to those that were done through mail (p's > .05). Based on initial STarT subgroup status, there were no differences in number of visits, follow-up rates, methods of follow-up, or current physical therapy status at 4-weeks (p's > .05).

Specific Aim 1

Construct Validity of the STarT Back Tool Classification Scheme at Initial Physical Therapy Evaluation

Baseline patient demographic, clinical, and psychological score characteristics are displayed in Table 5-1. As expected, the distribution of symptom duration (in days) demonstrated significant positive skewness and kurtosis therefore were described duration with nonparametric statistics. The STarT Back Tool categorized 53 (36.3%) patients as low risk, 55 (37.7%) as medium risk, and 38 (26.0%) as high risk. There were no differences in demographic characteristics amongst STarT subgroups (p's > .05). A comparison of baseline clinical characteristics indicated that patients in the low

risk category were more likely to have experienced a gradual onset of symptoms and less likely to have experienced a traumatic onset of symptoms (χ^2 (2) = 25.88, p < .001).

Comparisons amongst all STarT risk subgroups at indicated a dose-response relationship with several psychological and clinical measures. Specifically, patients initially allocated to the STarT low risk subgroup had lower initial FABQ-PA, PCS, TSK-11, and PHQ-9 scores compared to those categorized as being medium or high risk (p's < .05). This dose-response relationship was also consistent when using the STarT medium or high risk group as a reference for comparisons. Similar results were indicated for pain intensity (NPRS) ratings (p's < .001) and LBP-related disability (RMDQ) scores (p's < .001). Comparisons amongst STarT low and high risk subgroups indicated that patients initially allocated to the STarT low risk subgroup had lower initial FABQ-W and STAI-T scores compared to those categorized as high risk (p's < .05). Similar results were indicated for LBP-related disability (ODQ) and physical impairment (PII) scores (p's < .01). In addition to the findings above, the only unique finding for patients initially allocated to the STarT medium risk subgroup were greater initial LBP-related disability (ODQ) scores compared to those categorized as being low risk (p < .001).

Discrimination of Psychological Factors amongst STarT Back Tool Classification Status at Initial Physical Therapy Evaluation

Discriminant function analysis run with simultaneous entry method with six predictors (FABQ-PA: Wilks' λ = .80, p < .001; FABQ-W: Wilks' λ = .95, p = .031; PCS: Wilks' λ = .78, p < .001; TSK-11: Wilks' λ = .81, p < .001; PHQ-9: Wilks' λ = .69, p < .001; STAI-T: Wilks' λ = .94, p = .011;) suggested that each predictor contributed uniquely to initial STarT subgroup allocation and resulted in two discriminant functions

which is expected with 3 STarT subgroups. The overall test of the two functions (i.e., functions 1 and 2) was significant (χ^2 (12) = 76.19, Wilks' λ = .57, $p < .001$) indicating that predictor scores were able to discriminate amongst the three STarT subgroups. The test for function 2 alone was not significant (χ^2 (5) = 0.73, Wilks' λ = .99, p = .981), indicating that that after function 1 is removed, significant discrimination did not remain, therefore results will be reported only for function 1. Function 1 accounted for 42% (canonical R = .65) of the total relationship between predictors and STarT subgroups. The first discriminant function accounted for 99% of the between-group variability of STarT subgroup status. The pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions, as well as the standardized canonical discriminant function coefficients (analogous to multiple regression beta weights) are provided in Table 5-2. When discriminant function analyses result in multiple functions, the first function is considered the most important.[166] Therefore, based on the standardized coefficients for the first discriminant function in Table 5-2, depressive symptoms (PHQ-9) demonstrated the strongest positive relationship with the discriminant function, whereas fear-avoidance beliefs about physical activity (FABQ-PA) and pain catastrophizing (PCS) demonstrated moderate positive relationships and kinesiophobia (TSK-11) demonstrated the weakest positive relationship. Fear-avoidance beliefs about work (FABQ-W) and trait anxiety (STAI-T) demonstrated weak negative relationships with the discriminant function. Together, the functions were able to correctly classify 55.6% of the 3 STarT subgroups (75.0% of the Low Risk, 30.8% of the Medium Risk, and 63.2% of the High Risk), which is higher than the expected rate

from chance (estimated at 33% assuming equal probability of STarT risk group assignment).

Clustering of Psychological Factors at Initial Physical Therapy Evaluation without considering STarT Back Tool Classification Status

Results from our exploratory cluster analysis investigating how psychological factors clustered at intake, without consideration for STarT subgroup status, indicated that a 3-cluster solution was appropriate (Figure 5-1). Z-score transformations for psychological measure scores were required for cluster analysis procedures; however raw scores are reported for descriptive purposes because they are more clinically interpretable. Inspection of all predictor z-scores indicated that absolute values did not exceed 4.0 (range = -2.54 to 3.20), suggesting the data did not contain extreme outliers.[166, 167] Inspection of agglomeration coefficients from a hierarchical agglomerative cluster analysis of six psychological measures revealed that the percent change was large (46.8%) between the 3 and 2-cluster solutions with relatively smaller changes in preceding steps, suggesting a 3-cluster solution is appropriate, which was further confirmed by visual inspection of plotted agglomeration coefficients.[164, 165] Cluster 1 was labeled “Low Psychological Distress” (n = 84, 59%) and was comprised of individuals that were associated with low individual FAM-based psychological measure scores. Cluster 2 was labeled “High Psychological Distress” (n = 38, 27%) and was comprised of individuals that were associated with high individual FAM-based psychological measure scores. Cluster 3 was labeled “High Psychological Distress with Low Fear-Avoidance Work Beliefs” (n = 20, 14%) and was comprised of individuals that were associated with high individual FAM-based psychological measure scores with the exception being FABQ-W.

Differences were detected between clusters for STarT subgroup distributions (Figure 5-2). Chi-square analyses revealed that the Low Psychological Distress cluster had proportionally more STarT low risk patients (53.6%) compared to High Psychological Distress (10.5%) or High Psychological Distress with Low Fear-Avoidance Work Beliefs (15.0%) clusters and less STarT high risk patients (8.3%) compared to High Psychological Distress (47.4%) or High Psychological Distress with Low Fear-Avoidance Work Beliefs (65.0%) clusters (χ^2 (4) = 45.70, p < .001).

There were no differences in demographic variables detected between clusters (p's > .05), however there was a potential trend for less patients with work-related LBP in the High Psychological Distress with Low Fear-Avoidance Work Beliefs cluster (p = .06). Differences in initial clinical measures were detected between clusters for pain intensity ratings (NPRS; p < .001), LBP-related disability (ODQ and RMDQ; p's < .001), and physical impairment (PII; p < .05) scores. Post-hoc comparisons with Bonferonni correction revealed that the Low Psychological Distress cluster was associated with lower NPRS, ODQ, RMDQ, and PII scores compared to the High Psychological Distress cluster (p's < .001) and lower ODQ and RMDQ scores compared to the High Psychological Distress with Low Fear-Avoidance Work Beliefs cluster (p's = .001). Differences between clusters were also detected for all psychological measure scores (i.e., STarT total, STarT psychosocial, FABQ-PA, FABQ-W, PCS, TSK-11, PHQ-9, and STAI-T) (p's < .001). Post-hoc comparisons with Bonferonni correction revealed that the Low Psychological Distress cluster was associated with lower psychological measure scores compared to other clusters (p's < .001), with the exception of FABQ-W scores

where the High Psychological Distress cluster was associated with higher FABQ-W scores compared to other clusters (p's < .001).

Specific Aim 2

Pearson correlation coefficients indicated that initial FAM-related psychological measures were significantly correlated with each other (r = 0.19 to 0.57, p < .05) and with initial STarT psychosocial subscale scores (r = 0.20 to 0.62, p < .05). Regression diagnostics indicated that multicollinearity was not a concern amongst predictor variables for all models according to published guidelines.[168] Variance inflation factors (VIF) were not substantially greater than 1.0 and were all below 10.0 (range: 1.10 to 2.90), and tolerance estimates were greater than 0.2 (range: 0.34 to 0.91). For all regression models, assumptions of homoscedasticity and linearity of standardized residuals were met, and independence of residuals was confirmed through the Durbin-Watson statistic (range: 1.77 to 2.09).[168]

Predictive Validity of the STarT Back Tool (4-Week Outcomes)

4-week pain intensity (NPRS) scores

The final regression model for 4-week pain intensity ratings is reported in Table 5-3. In the first block, baseline NPRS scores accounted for 33% (F (1,112) = 54.79, p < .001) of the variance in 4-week pain intensity ratings. In the second through fourth blocks, demographic and clinical variables, STarT psychosocial scores, and other psychological measures contributed an additional 2% to 3% of variance, however did not reach statistical significance (model change p's > .05). In the final model explaining 42% of the variance in 4-week pain intensity ratings (F (14,113) = 5.02, p < .001), baseline NPRS scores (β = .67, p < .001) and household income (β = -.21, p = .017) contributed unique variance.

4-week disability (ODQ) scores

The final regression model for 4-week ODQ scores is reported in Table 5-4. In the first block, baseline ODQ scores accounted for 48% ($F(1,112) = 104.48$, $p < .001$) of the variance in 4-week ODQ disability scores. In the second block, demographic and clinical variables contributed an additional 7% ($F(6,106) = 2.64$, $p = .020$) of variance in 4-week ODQ disability scores. In the third and fourth blocks, STarT psychosocial scores and psychological measure scores contributed an additional 0.4% and 3% of variance respectively, however did not reach statistical significance (model change p's > .05). In the final model explaining 58% of the variance in 4-week ODQ scores ($F(14,113) = 9.89$, $p < .001$), baseline ODQ scores ($\beta = .57$, $p < .001$), age ($\beta = .19$, $p = .017$), and household income ($\beta = -.22$, $p = .004$) scores contributed unique variance and a potential trend was indicated for FABQ-PA scores ($\beta = .16$, $p = .065$).

4-week disability (RMDQ) scores

The final regression model for 4-week RMDQ scores is reported in Table 5-5. In the first block, baseline RMDQ scores accounted for 42% ($F(1,110) = 81.12$, $p < .001$) of the variance in 4-week RMDQ disability scores. In the second through fourth blocks, demographic and clinical variables, STarT psychosocial scores, and other psychological measures contributed an additional 1% to 4% of variance, however did not reach statistical significance (model change p's > .05). In the final model explaining 53% of the variance in 4-week RMDQ scores ($F(14,111) = 7.84$, $p < .001$), baseline RMDQ scores ($\beta = .61$, $p < .001$), age ($\beta = .26$, $p = .003$), household income ($\beta = -.18$, $p = .026$), and FABQ-PA scores ($\beta = .20$, $p = .027$) contributed unique variance.

4-week physical impairment (PII) scores

The final regression model for 4-week physical impairment scores is reported in Table 5-6. In the first and second blocks, baseline PII scores accounted for 41% ($F(1,95) = 66.44$, $p < .001$) and demographic and clinical variables contributed an additional 8% ($F(6,89) = 2.23$, $p = .047$) of variance in 4-week physical impairment scores. In the third block, STarT psychosocial scores contributed an additional 0.4% of variance, however did not reach statistical significance (model change $p > .05$). In the fourth block, other psychological measures contributed an additional 8% ($F(6,82) = 2.68$, $p = .020$) of variance in 4-week physical impairment scores. In the final model explaining 58% of the variance in 4-week PII scores ($F(14,96) = 7.95$, $p < .001$), baseline PII scores ($\beta = .51$, $p < .001$), age ($\beta = .22$, $p = .015$), and FABQ-W scores ($\beta = .25$, $p = .006$) contributed unique variance.

Predictive Validity of the STarT Back Tool (4-Week Outcome Change Scores)

4-week pain intensity (NPRS) raw change scores

The final regression model for intake to 4-week pain intensity raw change scores is reported in Table 5-7. In the first block, demographic and clinical variables accounted for 5% of the variance in change scores, however did not reach statistical significance ($p = .427$). In the second block, STarT psychosocial scores contributed an additional 10% ($F(1,106) = 11.92$, $p < .001$) of variance in change scores. In the third block, other psychological measures contributed an additional 4% of variance, however did not reach statistical significance (model change $p = .557$). In the final model explaining 19% of the variance in intake to 4-week pain intensity change scores ($F(13,113) = 1.79$, $p = .055$), household income ($\beta = .23$, $p = .023$) and STarT psychosocial scores ($\beta = .32$, $p = .027$) contributed unique variance.

4-week disability (ODQ) raw change scores

The final regression model for intake to 4-week ODQ raw change scores is reported in Table 5-8. In the first block, demographic and clinical variables accounted for 11% of the variance in change scores, however did not reach statistical significance ($p = .058$). In the second and third blocks, STarT psychosocial scores and psychological measure scores contributed an additional 2% and 4% of variance respectively, however did not reach statistical significance (model change p's > .05). In the final model explaining 16% of the variance in intake to 4-week ODQ change scores ($F (13,113) = 1.49$, $p = .133$), only household income ($\beta = .27$, $p = .011$) contributed unique variance.

4-week disability (RMDQ) raw change scores

The final regression model for intake to 4-week RMDQ raw change scores is reported in Table 5-9. In the first block, demographic and clinical variables accounted for 6% of the variance in change scores, however did not reach statistical significance ($p = .354$). In the second block, STarT psychosocial scores contributed an additional 14% ($F (1,104) = 17.67$, $p < .001$) of variance in change scores. In the third block, other psychological measures contributed an additional 4%, however did not reach statistical significance (model change $p = .451$). In the final model explaining 24% of the variance in intake to 4-week RMDQ change scores ($F(13,111) = 2.40$, $p = .007$), age ($\beta = -.22$, $p = .031$), household income ($\beta = .20$, $p = .042$), and STarT psychosocial scores ($\beta = .44$, $p = .002$), contributed unique variance.

4-week physical impairment (PII) raw change scores

The final regression model for intake to 4-week PII raw change scores is reported in Table 5-10. In the first block, demographic and clinical variables accounted for 8% of the variance in change scores, however did not reach statistical significance ($p = .280$).

In the second and third blocks, STarT psychosocial scores and psychological measure scores contributed an additional 0% and 10% of variance respectively, however did not reach statistical significance (model change p's > .05). In the final model explaining 18% of the variance in intake to 4-week PII change scores ($F(13,96) = 1.38$, $p = .189$), only the number of physical therapy visits ($\beta = .26$, $p = .041$), contributed unique variance.

Specific Aim 3

Prediction of Sustained STarT High Risk Allocation

At baseline and 4-weeks, STarT subgroup status was available for 82.2% (n = 120) of the total study sample. At intake, 27.5% (n = 33) of these patients were allocated to the STarT high risk subgroup in comparison to 6.7% (n = 8) at 4-weeks. Logistic regression analyses indicated that initial FABQ-PA scores were significant predictors of sustained STarT subgroup status at 4-weeks ($B = .28$, $p = .032$, $OR = 1.32$), with higher FABQ-PA scores associated with higher probability of being allocated to the STarT high risk subgroup (Table 5-11). Based on these results and specific to this study sample, each unit increase in FABQ-PA scores increases the odds of sustained STarT high risk subgroup status at 4-weeks from 1.00 to 1.32.

Sustained STarT High Risk Allocation and Outcomes

Results from an exploratory repeated measures ANOVA analysis run with only patients that were high risk at baseline (n = 33) indicated that there were no significant group (i.e., sustained or non-sustained high risk status) x time interactions for NPRS ($F(1,31) = 0.00$, $p = .999$, partial $\eta^2 = .00$) or RMDQ ($F(1,29) = 0.66$, $p = .422$, partial $\eta^2 = .02$) scores, with a potential trend for PII ($F(1,25) = 2.96$, $p = .098$, partial $\eta^2 = .11$) scores. However, there was a significant group x time interaction for ODQ ($F(1,31) =$

7.77, p = .009, partial η^2 = .20) scores. Patients allocated to STarT high risk at intake and changed to STarT low or medium risk at 4-weeks were associated with decreased ODQ scores (M = 11.3, sd = 13.5), while patients that remained in STarT high risk were associated with increased ODQ scores (M = 5.0, sd = 8.6).

Table 5-1. Baseline Patient Characteristics.

Variable	Total Sample (n = 146)	STarT Low Risk (n = 53)	STarT Medium Risk (n = 55)	STarT High Risk (n = 38)	p-value
Age (Years)	41.4 (13.5)	38.8 (14.0)	44.6 (12.8)	40.3 (13.0)	.069
Sex (Female)	89 (61.0%)	28 (52.8%)	37 (67.3%)	24 (63.2%)	.291
(Male)	57 (39.0%)	25 (47.2%)	18 (32.7%)	14 (36.8%)	
Race (Caucasian)	110 (75.3%)	39 (73.6%)	41 (74.5%)	30 (79.0%)	
(African American)	25 (17.1%)	12 (22.6%)	9 (16.4%)	4 (10.5%)	.635
(Other)	11 (7.6%)	2 (3.8%)	5 (9.1%)	4 (10.5%)	
Ethnicity (Hispanic or Latino)	13 (8.9%)	3 (5.7%)	6 (10.9%)	4 (10.5%)	
(Not Hispanic or Latino)	130 (89.0%)	49 (92.4%)	49 (89.1%)	32 (84.2%)	.579
(Missing)	3 (2.1%)	1 (1.9%)	0 (0.0%)	2 (5.3%)	
Employment status (Employed)	94 (64.4%)	35 (66.0%)	39 (70.9%)	20 (52.6%)	
(Unemployed)	39 (26.7%)	15 (28.3%)	9 (16.4%)	15 (39.5%)	.219
(Retired)	10 (6.8%)	2 (3.8%)	6 (10.9%)	2 (5.3%)	
(Missing)	3 (2.1%)	1 (1.9%)	1 (1.8%)	1 (2.6%)	
Education completed (Less than high school)	7 (4.8%)	3 (5.7%)	0 (0.0%)	4 (10.5%)	
(High school)	85 (58.2%)	30 (56.6%)	31 (56.4%)	24 (63.2%)	
(College)	43 (29.5%)	12 (22.6%)	21 (38.2%)	10 (26.3%)	.111
(Post-graduate)	10 (6.8%)	7 (13.2%)	3 (5.4%)	0 (0.0%)	
(Missing)	1 (0.7%)	1 (1.9%)	0 (0.0%)	0 (0.0%)	
Income level (< $20,000)	32 (21.9%)	15 (28.3%)	8 (14.5%)	9 (23.7%)	
($20,000 to $35,000)	22 (15.1%)	5 (9.4%)	12 (21.8%)	5 (13.2%)	
($35, 001 to $50,000)	20 (13.7%)	7 (13.2%)	10 (18.2%)	3 (7.9%)	.415
($50,001 to $70,000)	22 (15.1%)	6 (11.3%)	9 (16.4%)	7 (18.4%)	
(> $70,000)	45 (30.8%)	17 (32.1%)	15 (27.3%)	13 (34.2%)	
(Missing)	5 (3.4%)	3 (5.7%)	1 (1.8%)	1 (2.6%)	
Previous low back surgery (Yes)	26 (17.8%)	8 (15.1%)	11 (20.0%)	7 (18.4%)	.773
Symptom duration (Days)	483.6 (1157.7)	538.3 (1208.1)	455.2 (1234.3)	448.2 (991.7)	.913
Acute (≤ 14 days)	17 (11.6%)	3 (5.7%)	11 (20.0%)	3 (7.9%)	
Subacute (15-90 days)	56 (38.4%)	20 (37.7%)	17 (30.9%)	19 (50.0%)	.079
Chronic (≥ 91 days)	70 (47.9%)	29 (54.7%)	25 (45.5%)	16 (42.1%)	
(Missing)	3 (2.1%)	1 (1.9%)	2 (3.6%)	0 (0.0%)	

Table 5-1. Continued.

Variable	Total Sample (n = 146)	STarT Low Risk (n = 53)	STarT Medium Risk (n = 55)	STarT High Risk (n = 38)	p-value
Symptom onset (Gradual)	70 (47.9%)	34 (64.2%)	23 (41.8%)	13 (34.2%)	
(Sudden)	53 (36.3%)	13 (24.5%)	22 (40.0%)	18 (47.4%)	.042
(Traumatic)	21 (14.4%)	5 (9.4%)	10 (18.2%)	6 (15.8%)	
(Missing)	2 (1.4%)	1 (1.9%)	0 (0.0%)	1 (2.6%)	
Symptom location (Low back only)	49 (33.6%)	25 (47.2%)	15 (27.3%)	9 (23.7%)	
(Low back & buttock or thigh)	72 (49.3%)	23 (43.4%)	30 (54.5%)	19 (50.0%)	.059
(Low back & lower leg)	25 (17.1%)	5 (9.4%)	10 (18.2%)	10 (26.3%)	
Work-related LBP (Yes)	19 (13.0%)	8 (15.1%)	7 (12.7%)	4 (10.5%)	.813
FAM-Based Psychological Measures					
FABQ-PA (potential range, 0-24)[a]	14.6 (5.8)	11.8 (5.5)	14.6 (5.4)	18.4 (4.4)	<.001
FABQ-W (potential range, 0-42)[b]	12.8 (11.1)	9.6 (9.6)	13.7 (10.7)	15.8 (12.6)	.023
PCS (potential range, 0-52)[a]	16.9 (12.2)	10.7 (9.6)	17.1 (10.5)	25.3 (12.9)	<.001
TSK-11 (potential range, 11-44)[a]	25.2 (6.9)	21.9 (6.4)	25.2 (5.4)	29.7 (6.9)	<.001
PHQ-9 (potential range, 0-27)[a]	7.4 (6.1)	3.7 (3.8)	7.6 (5.5)	12.4 (6.0)	<.001
STAI-T (potential range, 20-80)[b]	36.0 (9.2)	33.4 (7.6)	36.0 (9.9)	39.5 (9.4)	.008
STarT Measures					
STarT – Overall score (potential range, 0-9)[a]	4.5 (2.5)	1.8 (1.0)	5.1 (1.1)	7.4 (1.1)	<.001
STarT – Psych score (potential range, 0-5)[a]	2.4 (1.6)	0.9 (0.8)	2.3 (0.9)	4.5 (0.5)	<.001
Clinical Outcome Measures					
NRS – Average (potential range, 0-10)[a]	5.4 (2.0)	4.4 (1.9)	5.6 (1.9)	6.6 (1.5)	<.001
ODQ (potential range, 0-100)[b]	32.5 (16.7)	20.0 (12.7)	37.0 (14.1)	43.5 (14.2)	<.001
RMDQ (potential range, 0-24)[a]	11.2 (6.0)	6.7 (4.5)	12.1 (5.1)	16.2 (4.2)	<.001
PII (potential range, 0-7)[b]	3.9 (1.8)	3.3 (1.6)	4.0 (2.0)	4.6 (1.6)	.006

All values represent means (standard deviations) or frequency counts (percentages).
[a] indicates (low risk < medium risk < high risk) ($p < .05$); [b] indicates (low risk < high risk) ($p < .05$).

Table 5-2. Coefficients of Psychological Measure Predictor Variables of the Discriminant Function.

	Discriminant Function 1	
FAM-Based Psychological Measure	**Standardized Coefficient [a]**	**Correlation Coefficient [b]**
FABQ-PA	0.449	0.579
FABQ-W	-0.126	0.256
PCS	0.351	0.621
TSK-11	0.083	0.567
PHQ-9	0.692	0.785
STAI-T	-0.119	0.300

Key: [a] indicates standardized canonical discriminant function coefficients; [b] indicates pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions.

Table 5-3. Four-Week Pain Intensity Scores (NPRS).

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Baseline NPRS score	.33	.32	.33	<.001
2. Demographic & Clinical	.37	.33	.04	.320
3. STarT Psychosocial score	.39	.34	.02	.067
4. Psychological measure scores	.42	.33	.03	.655

Final Model for 4-Week Pain Intensity (NPRS) Ratings (n = 114)

Variable	B	β	p-value	VIF
Baseline NPRS score	.81	.67	< .001	1.92
Age	.01	.07	.419	1.38
Sex	-.37	-.08	.341	1.10
Household income	-.31	-.21	.017	1.27
Surgery	.17	.03	.731	1.16
Duration of current symptoms	-3.09E-5	-.02	.838	1.11
PT visits	.003	.003	.975	1.39
STarT Psychosocial score	-.30	-.21	.104	2.76
FABQ-PA	.05	.12	.213	1.65
FABQ-W	.01	.07	.439	1.37
PCS	-.01	-.05	.655	2.50
TSK-11	.02	.05	.705	2.51
PHQ-9	-.02	-.05	.670	2.10
STAI-T	-.004	-.01	.878	1.52

Final model accounted for 42% of the variance in NPRS scores [$F_{(14,113)} = 5.02$, $p < .001$].

Table 5-4. Four-Week Disability Scores (ODQ).

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Baseline ODQ score	.48	.48	.48	< .001
2. Demographic & Clinical	.55	.52	.07	.020
3. STarT Psychosocial score	.55	.52	.004	.365
4. Psychological measure scores	.58	.52	.03	.327

Final Model for 4-Week Disability (ODQ) Scores (n = 114)

Variable	B	β	p-value	VIF
Baseline ODQ score	.59	.57	< .001	2.01
Age	.24	.19	.017	1.43
Sex	-3.68	-.10	.141	1.12
Household income	-2.39	-.22	.004	1.27
Surgery	1.13	.03	.720	1.23
Duration of current symptoms	.00	.02	.766	1.11
PT visits	-.68	-.10	.186	1.33
STarT Psychosocial score	.08	.01	.945	2.58
FABQ-PA	.48	.16	.065	1.66
FABQ-W	.08	.06	.474	1.41
PCS	-.004	-.003	.979	2.49
TSK-11	.02	.009	.932	2.51
PHQ-9	.16	.06	.548	2.10
STAI-T	-.13	-.07	.404	1.48

Final model accounted for 58% of the variance in ODQ scores [$F_{(14,113)} = 9.89$, $p < .001$].

Table 5-5. Four-Week Disability Scores (RMDQ).

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Baseline RMDQ score	.42	.42	.42	< .001
2. Demographic & Clinical	.48	.45	.06	.099
3. STarT Psychosocial score	.49	.45	.01	.306
4. Psychological measure scores	.53	.46	.04	.159

Final Model for 4-Week Disability (RMDQ) Scores (n = 112)

Variable	B	β	p-value	VIF
Baseline RMDQ score	.59	.61	< .001	2.24
Age	.11	.26	.003	1.41
Sex	.03	.002	.974	1.13
Household income	-.67	-.18	.026	1.29
Surgery	-.16	-.01	.889	1.21
Duration of current symptoms	-5.96E-6	-.001	.987	1.16
PT visits	-.22	-.10	.240	1.34
STarT Psychosocial score	-.73	-.20	.095	2.90
FABQ-PA	.21	.20	.027	1.71
FABQ-W	.02	.03	.685	1.42
PCS	.08	.17	.114	2.39
TSK-11	-.06	-.07	.532	2.61
PHQ-9	.001	.001	.989	2.20
STAI-T	-.04	-.06	.489	1.50

Final model accounted for 53% of the variance in RMDQ scores [$F_{(14,111)} = 7.84$, $p < .001$].

Table 5-6. Four-Week Physical Impairment Scores (PII).

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Baseline PII score	.41	.41	.41	< .001
2. Demographic & Clinical	.49	.45	.08	.047
3. STarT Psychosocial score	.49	.45	.004	.401
4. Psychological measure scores	.58	.50	.08	.020

Final Model for 4-Week Physical Impairment (PII) Scores (n = 97)

Variable	B	β	p-value	VIF
Baseline PII score	.54	.51	< .001	1.39
Age	.03	.22	.015	1.46
Sex	.17	.04	.587	1.15
Household income	-.17	-.13	.094	1.24
Surgery	.75	.15	.073	1.32
Duration of current symptoms	-1.38E-4	-.08	.307	1.12
PT visits	-.11	-.14	.121	1.59
STarT Psychosocial score	-.08	-.07	.536	2.38
FABQ-PA	.03	.08	.362	1.65
FABQ-W	.04	.25	.006	1.47
PCS	.005	.04	.740	2.30
TSK-11	.008	.03	.808	2.53
PHQ-9	-.008	-.03	.800	2.29
STAI-T	.01	.07	.480	1.81

Final model accounted for 58% of the variance in PII scores [$F_{(14,96)} = 7.95$, $p < .001$].

Table 5-7. Intake to 4-week Pain Intensity (NPRS) Change Scores.

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Demographic & Clinical	.05	.00	.05	.427
2. STarT Psychosocial score	.15	.09	.10	.001
3. Psychological measure scores	.19	.08	.04	.557

Final Model for Intake to 4-Week Pain Intensity (NPRS) Change Scores (n = 114)

Variable	B	β	p-value	VIF
Age	-.01	-.05	.603	1.33
Sex	.42	.10	.290	1.10
Household income	.29	.23	.023	1.27
Surgery	-.18	-.03	.717	1.17
Duration of current symptoms	1.90E-5	.01	.900	1.11
PT visits	-.03	-.04	.729	1.33
STarT Psychosocial score	.39	.32	.027	2.45
FABQ-PA	-.06	-.16	.179	1.64
FABQ-W	-.009	-.05	.626	1.31
PCS	.02	.12	.393	2.33
TSK-11	-.03	-.09	.530	2.44
PHQ-9	.03	.09	.510	2.05
STAI-T	-1.04E-4	-4.82E-4	.996	1.50

Final model accounted for 19% of the variance in NRS change scores [$F_{(13,113)} = 1.79$, $p = .055$].

Table 5-8. Intake to 4-Week Disability (ODQ) Change Scores.

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Demographic & Clinical	.11	.06	.11	.058
2. STarT Psychosocial score	.12	.07	.02	.137
3. Psychological measure scores	.16	.05	.04	.604

Final Model for Intake to 4-Week Disability (ODQ) Change Scores (n = 114)

Variable	B	β	p-value	VIF
Age	-.14	-.14	.190	1.35
Sex	5.16	.19	.056	1.10
Household income	2.24	.27	.011	1.27
Surgery	1.73	.05	.605	1.18
Duration of current symptoms	-1.13E-4	-.01	.913	1.11
PT visits	.66	.12	.240	1.33
STarT Psychosocial score	1.06	.13	.363	2.43
FABQ-PA	-.34	-.14	.219	1.64
FABQ-W	.05	.04	.664	1.31
PCS	.14	.14	.325	2.35
TSK-11	-.22	-.11	.436	2.44
PHQ-9	.03	.02	.903	2.04
STAI-T	.11	.07	.503	1.48

Final model accounted for 16% of the variance in ODQ change scores [$F_{(13,113)} = 1.49$, $p = .133$].

Table 5-9. Intake to 4-Week Disability (RMDQ) Change Scores.

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Demographic & Clinical	.06	.01	.06	.354
2. STarT Psychosocial score	.20	.14	.14	< .001
3. Psychological measure scores	.24	.14	.04	.451

Final Model for Intake to 4-Week Disability (RMDQ) Change Scores (n = 112)

Variable	B	β	p-value	VIF
Age	-.08	-.22	.031	1.36
Sex	.10	.01	.915	1.12
Household income	.65	.20	.042	1.29
Surgery	1.01	.08	.418	1.17
Duration of current symptoms	-2.02E-4	-.05	.602	1.14
PT visits	.16	.08	.436	1.33
STarT Psychosocial score	1.37	.44	.002	2.51
FABQ-PA	-.15	-.17	.133	1.67
FABQ-W	.03	.06	.536	1.33
PCS	-.07	-.18	.195	2.38
TSK-11	.02	.03	.846	2.58
PHQ-9	.08	.10	.413	2.09
STAI-T	.04	.07	.541	1.50

Final model accounted for 24% of the variance in RMD change scores [$F_{(13,111)} = 2.40$, $p = .007$].

Table 5-10. Intake to 4-Week Physical Impairment (PII) Change Scores.

Model	R^2	Adj. R^2	R^2 change	Model change (p-value)
1. Demographic & Clinical	.08	.02	.08	.280
2. STarT Psychosocial score	.08	.01	.00	.955
3. Psychological measure scores	.18	.05	.10	.139

Final Model for Intake to 4-Week Physical Impairment (PII) Change Scores (n = 97)

Variable	B	β	p-value	VIF
Age	-.01	-.09	.424	1.31
Sex	-.03	-.01	.927	1.14
Household income	.18	.17	.119	1.24
Surgery	-.40	-.10	.384	1.29
Duration of current symptoms	8.29E-5	.06	.589	1.11
PT visits	.16	.26	.041	1.56
STarT Psychosocial score	.18	.19	.226	2.33
FABQ-PA	-.01	-.05	.684	1.64
FABQ-W	-.03	-.21	.076	1.44
PCS	-.01	-.08	.600	2.29
TSK-11	-.04	-.16	.313	2.45
PHQ-9	.01	.05	.758	2.29
STAI-T	-.004	-.03	.831	1.80

Final model accounted for 18% of the variance in PII change scores [$F_{(13,96)} = 1.38$, $p = .189$].

Table 5-11. Overall Logistic Regression Model Predicting STarT High Risk Subgroup Status at 4-weeks.

Predictor	B	S.E. B	Wald's χ^2 (df = 1)	p-value	Exp(B) (95% CI)
Age	.08	.05	2.28	.131	1.08 (0.98 – 1.19)
Surgery	-19.56	7017.66	.00	.998	0.00 (---)
FABQ-PA	.28	.13	4.59	.032	1.32 (1.02 – 1.70)
PCS	.03	.06	.22	.639	1.03 (0.92 – 1.15)
TSK-11	-.01	.11	.01	.921	0.99 (0.80 – 1.22)
PHQ-9	.05	.09	.32	.569	1.05 (0.89 – 1.25)
STAI-T	.08	.05	2.71	.100	1.08 (0.98 – 1.19)
Constant	-14.74	4.89	9.08	.002	---

Block 1 (χ^2 (2) = 6.73, p = .034); Block 2 (χ^2 (7) = 22.52, p = .002); Goodness-of-fit test: Hosmer & Lemeshow (χ^2 (8) = 2.65, p = .955); R^2-type indices: Hosmer & Lemeshow R^2 = .384 (equation = model χ^2 / null model -2LL), Cox & Snell R^2 = .172, Nagelkerke R^2 = .443.

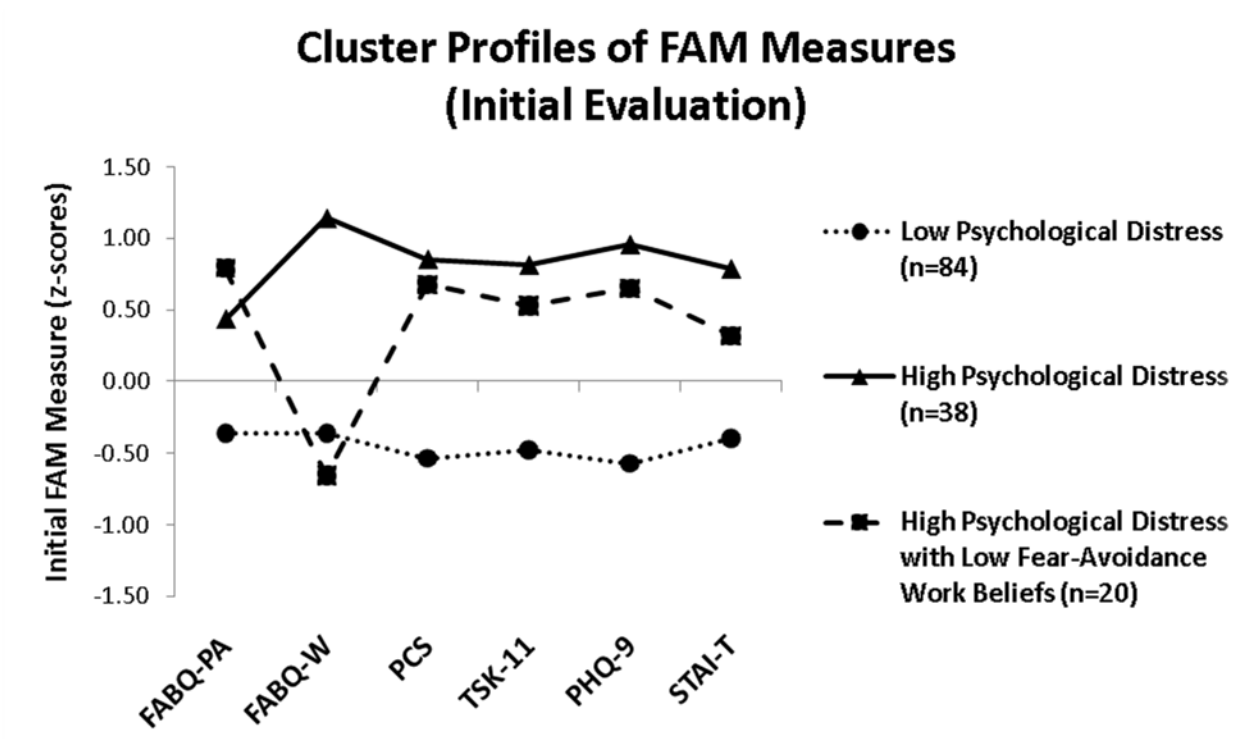

Figure 5-1. Cluster profiles using initial FAM measure scores.

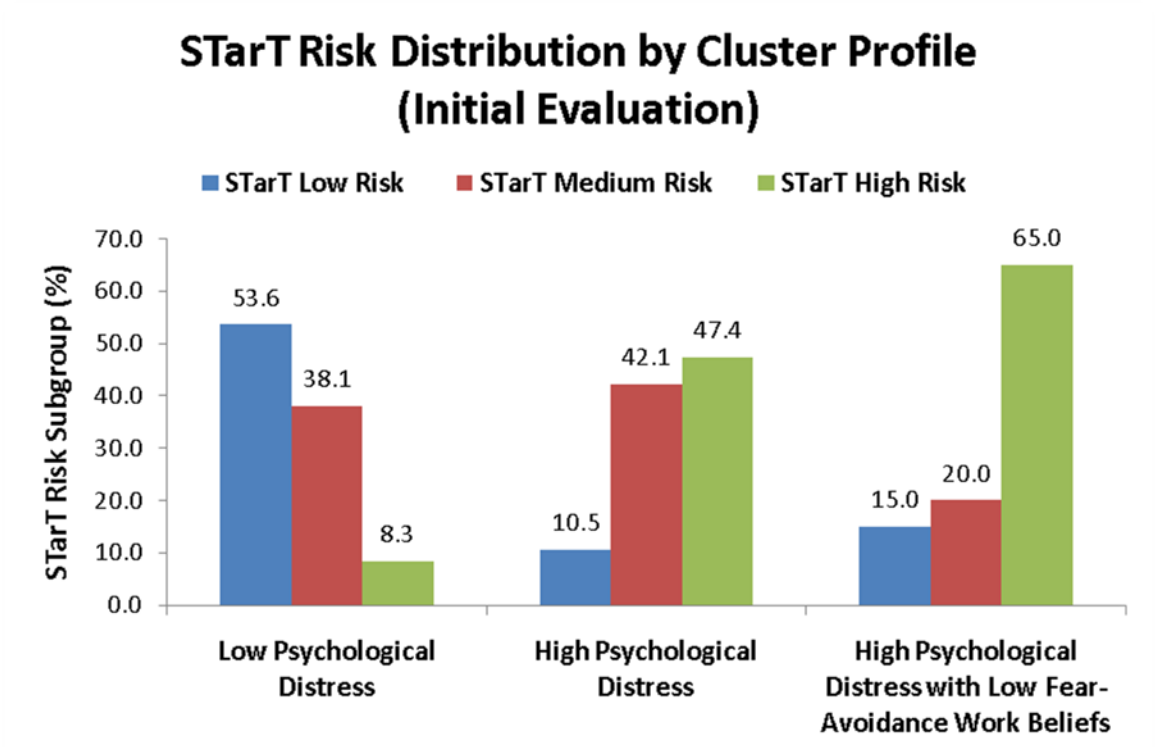

Figure 5-2. STarT risk distribution amongst cluster profiles.

CHAPTER 6
DISCUSSION

In the profession of physical therapy, "Psychologically Informed Practice" has recently been presented as a secondary prevention approach for chronic LBP that emphasizes routine and specific identification of potentially modifiable psychological risk factors.[13] Validation of measurement instruments in physical therapy settings has been highlighted as a top research priority before recommending their use in clinical practice.[13] The STarT Back Tool has been developed and intended for use in primary care settings for subgrouping LBP patients at-risk for poor future disability.[120] Utilization of the STarT Back Tool in outpatient physical therapy settings is promising,[121] however has not been extensively investigated. Therefore, the purpose of this dissertation was to: 1) determine the validity of the STarT Back Tool in outpatient physical therapy settings and 2) test the clinical utility of the STarT Back Tool in comparison to commonly used individual Fear-Avoidance Model psychological screening measures. We sought to achieve this purpose through three specific and several supporting exploratory aims.

Statement of Principal Findings

The STarT Back Tool classification scheme demonstrated good construct validity based on its ability to discriminate patients on initial clinical and psychological measures during the physical therapy evaluation. Most individual Fear-Avoidance Model psychological measures demonstrated moderate to strong abilities in discriminating initial STarT subgroup status and similar psychological profiles were generated when using cluster analysis to develop independent psychological subgroups. Both STarT psychosocial subscale and individual psychological measures demonstrated poor predictive validity for 4-week clinical outcomes, after controlling for initial scores.

However, our findings suggested that STarT psychosocial subscale scores may be important for predicting raw pain intensity and disability change scores. Finally, fear-avoidance beliefs about physical activity were important for predicting continued STarT high risk categorization after 4-weeks of physical therapy. This finding has potential clinical implications for the STarT Back Tool as a treatment monitoring measure because patients that remained high risk at 4-weeks reported worsening disability (i.e., higher ODQ scores) from intake while patients that changed STarT categorization from high risk to low or medium risk reported improved disability over 4 weeks. Future analyses of data from this study will determine if those that sustained STarT high risk status at 4 weeks also had poorer 6 month outcomes.

Strengths and Weaknesses of Study

This study consisted of several strengths. First, this was only the second study[121] we are aware of that has incorporated STarT Back Tool data from physical therapy settings. This may have important implications because the STarT Back Tool was developed and intended for use in primary care settings.[120, 122, 124, 128, 169] Second, in an attempt to meet previous research priorities,[5, 11, 13, 56, 126] the STarT Back Tool was compared to individual Fear-Avoidance Model psychological measures using a study design with standardized timing of follow-up assessments and measurement of multiple outcome domains. Finally, this study focused on potentially modifiable psychological risk factors which have important implications because future studies involving matched treatment approaches may consider incorporating the results of this study.

This study also consisted of several limitations. First, we did not measure the influence that other potentially important psychological factors (e.g., self-efficacy, preference, expectation) have on LBP outcomes.[116, 170] Second, physical therapy

treatment was not standardized in this study. Specifically, clinicians were not required to review psychological measure scores and treatment was not tailored to address psychological factors which may have influenced our results. Finally, we did not have access to 6-month outcomes during preparation of this dissertation which have potential to further support the findings of this study.

Comparison to Other Studies

The findings of this study are similar to previous studies that have used the STarT Back Tool. Hill and colleagues[120] indicated that initial STarT Back 'overall' and 'psychosocial' scores best discriminated initial self-reported disability and psychosocial reference standards respectively in primary care settings. Importantly, in that study, reference standards were based on dichotomized cutoff scores which may be associated with clinical limitations[13] and other outcome domains (e.g., pain intensity, physical impairment) were not included in validity analyses. Fritz and colleagues[121] reported similar findings to Hill and colleagues[120] for pain intensity and disability when measured on continuous scales in outpatient physical therapy settings, however did not include individual psychological measures. Therefore, a potential strength of this study in comparison to previous studies involving the STarT Back Tool includes the use of multiple outcome domains and several individual Fear-Avoidance Model measures. The results of this study are similar to Fritz and colleagues[121] based on relationships between initial STarT subgroup status and initial pain and disability, however also provides additional information involving initial physical impairment. Potentially, this has important clinical implications for physical therapists in outpatient physical therapy settings based on the STarT Back subgrouping scheme and its ability to discriminate among multiple outcome domains.

Foster and colleagues found that after controlling for demographic and baseline clinical characteristics, only 11 out of 20 psychological factors investigated accounted for between 0.5% to 4.9% ($p < .01$) additional variance in LBP-related disability at 6-months in primary care settings.[116] Our study findings are similar based on the addition of individual psychological measures or STarT psychosocial scores accounting for between 0.4% to 4.0% additional variance in pain intensity and disability at 4-weeks, however none reached statistical significance ($p > .05$). Our study findings potentially expand on Foster and colleagues[116] based on the addition of individual psychological measures accounting for 8.0% additional variance in physical impairment at 4-weeks ($p < .05$). Our study also provided alternative regression modeling techniques to investigate for differences based on using outcome change scores as dependent variables. These results indicated similar results based on the addition of individual psychological measures accounting for between 4.0% to 10.0% ($p > .05$) additional variance in all outcomes at 4-weeks. In contrast, STarT psychosocial scores accounted for 10.0% and 14.0% ($p < .001$) additional variance in pain intensity and disability (i.e., RMDQ) respectively.

With the exception of fear-avoidance beliefs about work for predicting physical impairment at 4-weeks, fear-avoidance beliefs about physical activity was the only other important individual psychological factor as it predicted both disability and patients that remained STarT high risk at 4-weeks. Similar to the results of this study, the influence of initial fear-avoidance beliefs about physical activity on disability following 4-weeks of physical therapy has been implicated in previous studies by our group.[78, 79] When attempting to interpret the relationship between fear-avoidance beliefs about work and

outcomes, one potential imitation to this study is that only 13.0% of patients reported work-related LBP. Previous studies that have reported on the influence of fear-avoidance work beliefs have primarily consisted of patients with work-related LBP or appropriate numbers to conduct subgroup analyses.[84, 111] Furthermore, it has been suggested that fear may be a more important predictor of chronic LBP outcomes and may not be as important during earlier stages of LBP.[91]

Potential Implications

A small portion of patients with acute and chronic LBP seek physical therapy services[41, 42] and previous findings indicate that a small percentage of patients with chronic LBP are utilizing a majority of healthcare resources associated with treatment for LBP.[44, 45] Considering that potentially modifiable risk factors have been identified with increased physical therapy utilization for LBP, a standardized screening process may assist in clinical decision-making processes. For example, identifying patients at high-risk for developing chronic LBP may influence treatment approaches in the form of tailored physical therapy interventions, psychological counseling, or both. Realistically, such a screening process would require a dynamic approach including consistent monitoring of psychological symptoms and response to treatment over an episode of physical therapy care, as opposed to a “one size fits all” intervention approach based solely on intake measures. Specifically, a dynamic screening process may consist of more information than is obtained during the initial evaluation as previous suggestions have indicated that some psychological factors (e.g., fear-avoidance beliefs) may be more important as prognostic indicators for chronic LBP, while others (e.g., depressive symptoms) may be more important at earlier stages.[91]

Unanswered Questions and Future Research

The American Pain Society and American College of Physicians have recommended routine assessment of psychological risk factors for patients with LBP.[171] However, a previous review study by our group[172] has suggested that psychological factors are not adequately assessed during development of clinical decision tools for physical therapy interventions. Future studies that use subgrouping paradigms to match patients with appropriate physical therapy interventions have potential to benefit from incorporating a standardized dynamic psychological screening procedure consisting of both individual Fear-Avoidance Model based measures and the STarT Back Screening Tool. As previously mentioned, future analyses of 6-month data may provide additional information as to the implications of initial psychological screening information or if continual monitoring over an episode of care is more appropriate. Therefore, a potentially interesting future study may consist of integrating two separate subgrouping paradigms for the management of LBP in the physical therapy setting. For example, the Treatment-Based Classification System[173] may provide useful information for initial treatment approaches and subsequent decisions to incorporate supplemental behavioral interventions could be based on treatment monitoring via STarT Back Tool risk allocation, but tailored to match responses from individual Fear-Avoidance Model measures.

Recently, a clinical trial comparing a stratified management care approach for LBP (based on STarT Back subgroup allocation) to current best practice was conducted in the primary care setting.[174] Results of that study indicated that after adjusting for confounding factors, changes in disability were significantly higher for the stratified care group compared to the current best practice group at 4-months (effect size = 0.32) and

12-months (effect size = 0.19). Moreover, stratified care was associated with a mean increase in generic health benefit and cost-savings at 12-months.[174] While the findings of this study have potential implications in the primary care, they require further testing before being translated to physical therapy settings.

Finally, this study investigated the influence of initial psychological measures on outcomes following 4-weeks of physical therapy. Future studies should investigate if changes in individual psychological measures are indicative of improved outcomes; and if so – which measures. Previous studies have indicated that changes in pain catastrophizing,[175, 176] fear-avoidance beliefs,[103] or depressive symptoms[177] were predictive of future outcomes, potentially suggesting that changes in psychological measures may provide more clinically important information than initial measures alone. Furthermore, future studies could extend the results of this study by investigating which changes in individual psychological measures are indicative of changes in STarT classification. Again, these questions may be answered when 6-month outcomes for this study are analyzed in the near future.

CHAPTER 7
CONCLUSIONS

This study investigated the validity and clinical utility of the STarT Back Screening Tool in comparison to several Fear-Avoidance Model based single-construct psychological measures in outpatient physical therapy settings. Results of this study indicated that The STarT Back Tool subgrouping scheme demonstrated good construct validity at initial physical therapy evaluation based on concurrent clinical and psychological measures. Both STarT psychosocial subscale and individual Fear-Avoidance Model measures demonstrated poor predictive validity for 4-week clinical outcomes, however our findings suggest that STarT psychosocial subscale scores may be important for predicting pain intensity and disability raw change scores. Finally, baseline fear-avoidance beliefs about physical activity were important for predicting those remaining as STarT high risk categorization at 4-weeks. This finding has potential clinical implications for the STarT Back Tool as a treatment monitoring measure because patients that remained high risk at 4-weeks reported worse disability from intake while patients that changed from high risk to low or medium risk reported improved disability. Based on the results of this study, immediate future research priorities include determining if initial or continual monitoring of psychological symptoms over an episode of care is more appropriate as a prognostic indicator for poor 6-month outcomes and if changes in individual psychological measures are indicative of changes in STarT classification or clinical outcomes. Further research priorities include development of appropriate psychological classification methods for the comprehensive management of LBP in physical therapy settings.

The primary purpose of this dissertation was to determine “optimal” screening procedures for psychological distress in patients seeking physical therapy for LBP. Based on these data, incorporating the STarT Back and FABQ-PA as part of a dynamic psychological screening process is recommended for use in outpatient physical therapy settings based on the following reasons: 1) the STarT Back demonstrated the ability to differentiate patients based on individual psychological and clinical characteristic measures at initial evaluation and 2) the FABQ-PA demonstrated the ability to detect which patients sustained STarT high risk status following 4-weeks of PT. As previously mentioned, future analysis of 6-month data potentially will either support or refute these recommendations and provide further insight as to the importance of screening for psychological distress at initial evaluation compared to monitoring of treatment responses over the episode of PT care.

LIST OF REFERENCES

1. From the Centers for Disease Control and Prevention. Prevalence of disabilities and associated health conditions among adults--United States, 1999. *JAMA*, 2001. 285(12): 1571-2.

2. CDC. *Prevalence and most common causes of disability among adults - United States, 2005.*, in *Morbidity and Mortality Weekly Report 58(16)*. 2009. p. 421-426.

3. Gureje O, Von Korff M, Simon GE, et al. Persistent pain and well-being: a World Health Organization Study in Primary Care. *JAMA*, 1998. 280(2): 147-51.

4. Deyo RA. Low-back pain. *Sci Am*, 1998. 279(2): 48-53.

5. Pransky G, Borkan JM, Young AE, et al. Are we making progress?: the tenth international forum for primary care research on low back pain. *Spine (Phila Pa 1976)*, 2011. 36(19): 1608-14.

6. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J*, 2008. 8(1): 8-20.

7. Jones GT, Johnson RE, Wiles NJ, et al. Predicting persistent disabling low back pain in general practice: a prospective cohort study. *Br J Gen Pract*, 2006. 56(526): 334-41.

8. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ*, 2006. 332(7555): 1430-4.

9. Hilfiker R, Bachmann LM, Heitz CA, et al. Value of predictive instruments to determine persisting restriction of function in patients with subacute non-specific low back pain. Systematic review. *Eur Spine J*, 2007. 16(11): 1755-75.

10. Blyth FM, Macfarlane GJ, Nicholas MK. The contribution of psychosocial factors to the development of chronic pain: the key to better outcomes for patients? *Pain*, 2007. 129(1-2): 8-11.

11. Nicholas MK, Linton SJ, Watson PJ, et al. Early identification and management of psychological risk factors ("yellow flags") in patients with low back pain: a reappraisal. *Phys Ther*, 2011. 91(5): 737-53.

12. Nicholas MK, George SZ. Psychologically informed interventions for low back pain: an update for physical therapists. *Phys Ther*, 2011. 91(5): 765-76.

13. Main CJ, George SZ. Psychologically informed practice for management of low back pain: future directions in practice and research. *Phys Ther*, 2011. 91(5): 820-4.

14. Szklo M, Nieto, FJ., *Measures of Disease Occurrence and Association*, in *Epidemiology: Beyond the Basics*. 2007, Jones and Bartlett: Sudbury, MA.

15. Stanton TR, Latimer J, Maher CG, et al. Definitions of recurrence of an episode of low back pain: a systematic review*. Spine (Phila Pa 1976)*, 2009. 34(9): E316-22.

16. Stanton TR, Latimer J, Maher CG, et al. A modified Delphi approach to standardize low back pain recurrence terminology*. Eur Spine J*, 2011. 20(5): 744-52.

17. Pengel LH, Herbert RD, Maher CG, et al. Acute low back pain: systematic review of its prognosis*. BMJ*, 2003. 327(7410): 323.

18. Stanton TR, Henschke N, Maher CG, et al. After an episode of acute low back pain, recurrence is unpredictable and not as common as previously thought*. Spine (Phila Pa 1976)*, 2008. 33(26): 2923-8.

19. Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998*. J Spinal Disord*, 2000. 13(3): 205-17.

20. Loney PL, Stratford PW. The prevalence of low back pain in adults: a methodological review of the literature*. Phys Ther*, 1999. 79(4): 384-96.

21. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002*. Spine (Phila Pa 1976)*, 2006. 31(23): 2724-7.

22. Strine TW, Hootman JM. US national prevalence and correlates of low back and neck pain among adults*. Arthritis Rheum*, 2007. 57(4): 656-65.

23. Hardt J, Jacobsen C, Goldberg J, et al. Prevalence of chronic pain in a representative sample in the United States*. Pain Med*, 2008. 9(7): 803-12.

24. Freburger JK, Holmes GM, Agans RP, et al. The rising prevalence of chronic low back pain*. Arch Intern Med*, 2009. 169(3): 251-8.

25. Deutscher D, Hart DL, Dickstein R, et al. Implementing an integrated electronic outcomes and electronic health record process to create a foundation for clinical practice improvement*. Phys Ther*, 2008. 88(2): 270-85.

26. Hart DL, Werneke MW, George SZ, et al. Screening for elevated levels of fear-avoidance beliefs regarding work or physical activities in people receiving outpatient therapy*. Phys Ther*, 2009. 89(8): 770-85.

27. Deutscher D, Horn SD, Dickstein R, et al. Associations between treatment processes, patient characteristics, and outcomes in outpatient physical therapy practice. *Arch Phys Med Rehabil*, 2009. 90(8): 1349-63.

28. Cherry DK, Burt CW, Woodwell DA. National Ambulatory Medical Care Survey: 2001 summary. *Adv Data*, 2003(337): 1-44.

29. Raofi S, Schappert SM. Medication therapy in ambulatory medical care: United States, 2003-04. *Vital Health Stat 13*, 2006(163): 1-40.

30. Deyo RA, Mirza SK, Turner JA, et al. Overtreating chronic back pain: time to back off? *J Am Board Fam Med*, 2009. 22(1): 62-8.

31. *Number of Events for Selected Conditions by Type of Service: United States.* 2008, Agency for Healthcare Research and Quality.

32. Soni A. *Back Problems: Use and Expenditures for the U.S. Adult Population, 2007*. 2010, Agency for Healthcare Research and Quality: Rockville, MD.

33. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*, 2008. 299(6): 656-64.

34. Luo X, Pietrobon R, Sun SX, et al. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine (Phila Pa 1976)*, 2004. 29(1): 79-86.

35. Nimgade A, McNeely E, Milton D, et al. Increased expenditures for other health conditions after an incident of low back pain. *Spine (Phila Pa 1976)*, 2010. 35(7): 769-77.

36. *National Institute of Neurological Disorders and Stroke. Low back pain fact sheet*. 2003; Available from: http://www.ninds.nih.gov/disorders/backpain/detail_backpain.htm.

37. Ricci JA, Stewart WF, Chee E, et al. Back pain exacerbations and lost productive time costs in United States workers. *Spine (Phila Pa 1976)*, 2006. 31(26): 3052-60.

38. Stewart WF, Ricci JA, Chee E, et al. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA*, 2003. 290(18): 2443-54.

39. Goetzel RZ, Hawkins K, Ozminkowski RJ, et al. The health and productivity cost burden of the "top 10" physical and mental health conditions affecting six large U.S. employers in 1999. *J Occup Environ Med*, 2003. 45(1): 5-14.

40. Guo HR, Tanaka S, Halperin WE, et al. Back pain prevalence in US industry and estimates of lost workdays. *Am J Public Health*, 1999. 89(7): 1029-35.

41. Mielenz TJ, Carey TS, Dyrek DA, et al. Physical therapy utilization by patients with acute low back pain. *Phys Ther*, 1997. 77(10): 1040-51.

42. Freburger JK, Carey TS, Holmes GM. Physical therapy for chronic low back pain in North Carolina: overuse, underuse, or misuse? *Phys Ther*, 2011. 91(4): 484-95.

43. Gellhorn AC, Chan L, Martin B, et al. Management Patterns in Acute Low Back Pain: The Role of Physical Therapy. *Spine (Phila Pa 1976)*, 2010.

44. Maetzel A, Li L. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Pract Res Clin Rheumatol*, 2002. 16(1): 23-30.

45. Asche CV, Kirkness CS, McAdam-Marx C, et al. The societal costs of low back pain: data published between 2001 and 2007. *J Pain Palliat Care Pharmacother*, 2007. 21(4): 25-33.

46. Lethem J, Slade PD, Troup JD, et al. Outline of a Fear-Avoidance Model of exaggerated pain perception--I. *Behav Res Ther*, 1983. 21(4): 401-8.

47. Vlaeyen JW, Kole-Snijders AM, Boeren RG, et al. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*, 1995. 62(3): 363-72.

48. Chaves JF, Brown JM. Spontaneous cognitive strategies for the control of clinical pain and stress. *J Behav Med*, 1987. 10(3): 263-76.

49. Goubert L, Crombez G, Van Damme S. The role of neuroticism, pain catastrophizing and pain-related fear in vigilance to pain: a structural equations approach. *Pain*, 2004. 107(3): 234-41.

50. Sullivan MJ, Thorn B, Haythornthwaite JA, et al. Theoretical perspectives on the relation between catastrophizing and pain. *Clin J Pain*, 2001. 17(1): 52-64.

51. Sullivan M, Bishop S, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychol Assess*, 1995. 7: 524-532.

52. Watson D, Clark LA, Carey G. Positive and negative affectivity and their relation to anxiety and depressive disorders. *J Abnorm Psychol*, 1988. 97(3): 346-53.

53. Watson D, Pennebaker JW. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychol Rev*, 1989. 96(2): 234-54.

54. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, 2000. 85(3): 317-32.

55. Leeuw M, Goossens ME, Linton SJ, et al. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. *J Behav Med*, 2007. 30(1): 77-94.

56. Pincus T, Smeets RJ, Simmonds MJ, et al. The fear avoidance model disentangled: improving the clinical utility of the fear avoidance model. *Clin J Pain*, 2010. 26(9): 739-46.

57. Hayden JA, Chou R, Hogg-Johnson S, et al. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*, 2009. 62(8): 781-796 e1.

58. Blyth FM, March LM, Brnabic AJ, et al. Chronic pain in Australia: a prevalence study. *Pain*, 2001. 89(2-3): 127-34.

59. Hancock M, Herbert RD, Maher CG. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Phys Ther*, 2009. 89(7): 698-704.

60. Kamper SJ, Maher CG, Hancock MJ, et al. Treatment-based subgroups of low back pain: a guide to appraisal of research studies and a summary of current evidence. *Best Pract Res Clin Rheumatol*, 2010. 24(2): 181-91.

61. Linton SJ. A review of psychological risk factors in back and neck pain. *Spine (Phila Pa 1976)*, 2000. 25(9): 1148-56.

62. Mannion AF, Dolan P, Adams MA. Psychological questionnaires: do "abnormal" scores precede or follow first-time low back pain? *Spine (Phila Pa 1976)*, 1996. 21(22): 2603-11.

63. Bigos SJ, Battie MC, Spengler DM, et al. A prospective study of work perceptions and psychosocial factors affecting the report of back injury. *Spine (Phila Pa 1976)*, 1991. 16(1): 1-6.

64. Linton SJ, Buer N, Vlaeyen J, et al. Are fear-avoidance beliefs related to a new episode of back pain? A prospective study. *Psychol Health*, 2000. 14(6): 1051-9.

65. Linton SJ. Do psychological factors increase the risk for back pain in the general population in both a cross-sectional and prospective analysis? *Eur J Pain*, 2005. 9(4): 355-61.

66. Van Nieuwenhuyse A, Somville PR, Crombez G, et al. The role of physical workload and pain related fear in the development of low back pain in young workers: evidence from the BelCoBack Study; results after one year of follow up. *Occup Environ Med*, 2006. 63(1): 45-52.

67. Picavet HS, Vlaeyen JW, Schouten JS. Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *Am J Epidemiol*, 2002. 156(11): 1028-34.

68. Croft PR, Papageorgiou AC, Ferry S, et al. Psychologic distress and low back pain. Evidence from a prospective study in the general population. *Spine (Phila Pa 1976)*, 1995. 20(24): 2731-7.

69. Leino P, Magni G. Depressive and distress symptoms as predictors of low back pain, neck-shoulder pain, and other musculoskeletal morbidity: a 10-year follow-up of metal industry employees. *Pain*, 1993. 53(1): 89-94.

70. Papageorgiou AC, Macfarlane GJ, Thomas E, et al. Psychosocial factors in the workplace--do they predict new episodes of low back pain? Evidence from the South Manchester Back Pain Study. *Spine (Phila Pa 1976)*, 1997. 22(10): 1137-42.

71. Peters ML, Vlaeyen JW, Weber WE. The joint contribution of physical pathology, pain-related fear and catastrophizing to chronic back pain disability. *Pain*, 2005. 113(1-2): 45-50.

72. Severeijns R, Vlaeyen JW, van den Hout MA, et al. Pain catastrophizing predicts pain intensity, disability, and psychological distress independent of the level of physical impairment. *Clin J Pain*, 2001. 17(2): 165-72.

73. Meyer K, Tschopp A, Sprott H, et al. Association between catastrophizing and self-rated pain and disability in patients with chronic low back pain. *J Rehabil Med*, 2009. 41(8): 620-5.

74. Crombez G, Vlaeyen JW, Heuts PH, et al. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*, 1999. 80(1-2): 329-39.

75. Grotle M, Vollestad NK, Veierod MB, et al. Fear-avoidance beliefs and distress in relation to disability in acute and chronic low back pain. *Pain*, 2004. 112(3): 343-52.

76. McCracken LM, Gross RT, Aikens J, et al. The assessment of anxiety and fear in persons with chronic pain: a comparison of instruments. *Behav Res Ther*, 1996. 34(11-12): 927-33.

77. Asmundson GJ, Norton GR. Anxiety sensitivity in patients with physically unexplained chronic back pain: a preliminary report. *Behav Res Ther*, 1995. 33(7): 771-7.

78. George SZ, Valencia C, Beneciuk JM. A psychometric investigation of fear-avoidance model measures in patients with chronic low back pain. *J Orthop Sports Phys Ther*, 2010. 40(4): 197-205.

79. George SZ, Calley D, Valencia C, et al. Clinical Investigation of Pain-related Fear and Pain Catastrophizing for Patients With Low Back Pain. *Clin J Pain*, 2011. 27(2): 108-15.

80. Sieben JM, Vlaeyen JW, Portegijs PJ, et al. A longitudinal study on the predictive validity of the fear-avoidance model in low back pain. *Pain*, 2005. 117(1-2): 162-70.

81. Swinkels-Meewisse IE, Roelofs J, Schouten EG, et al. Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. *Spine (Phila Pa 1976)*, 2006. 31(6): 658-64.

82. Sieben JM, Vlaeyen JW, Tuerlinckx S, et al. Pain-related fear in acute low back pain: the first two weeks of a new episode. *Eur J Pain*, 2002. 6(3): 229-37.

83. George SZ, Bialosky JE, Donald DA. The centralization phenomenon and fear-avoidance beliefs as prognostic factors for acute low back pain: a preliminary investigation involving patients classified for specific exercise. *J Orthop Sports Phys Ther*, 2005. 35(9): 580-8.

84. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain*, 2001. 94(1): 7-15.

85. Burton AK, McClune TD, Clarke RD, et al. Long-term follow-up of patients with low back pain attending for manipulative care: outcomes and predictors. *Man Ther*, 2004. 9(1): 30-5.

86. Boersma K, Linton SJ. Expectancy, fear and pain in the prediction of chronic pain and disability: a prospective analysis. *Eur J Pain*, 2006. 10(6): 551-7.

87. Klenerman L, Slade PD, Stanley IM, et al. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. *Spine (Phila Pa 1976)*, 1995. 20(4): 478-84.

88. Henschke N, Maher CG, Refshauge KM, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *BMJ*, 2008. 337: a171.

89. Koleck M, Mazaux JM, Rascle N, et al. Psycho-social factors and coping strategies as predictors of chronic evolution and quality of life in patients with low back pain: a prospective study. *Eur J Pain*, 2006. 10(1): 1-11.

90. Hadjistavropoulos HD, Asmundson GJ, Kowalyk KM. Measures of anxiety: is there a difference in their ability to predict functioning at three-month follow-up among pain patients? *Eur J Pain*, 2004. 8(1): 1-11.

91. Pincus T, Vogel S, Burton AK, et al. Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. *Arthritis Rheum*, 2006. 54(12): 3999-4010.

92. Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. *Man Ther*, 2008. 13(1): 12-28.

93. Hartvigsen J, Lings S, Leboeuf-Yde C, et al. Psychosocial factors at work in relation to low back pain and consequences of low back pain; a systematic, critical review of prospective cohort studies. *Occup Environ Med*, 2004. 61(1): e2.

94. Iles RA, Davidson M, Taylor NF. Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review. *Occup Environ Med*, 2008. 65(8): 507-17.

95. Mallen CD, Peat G, Thomas E, et al. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract*, 2007. 57(541): 655-61.

96. Main CJ, Wood PL, Hollis S, et al. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine (Phila Pa 1976)*, 1992. 17(1): 42-52.

97. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA*, 2010. 303(13): 1295-302.

98. Hill JC, Fritz JM. Psychosocial influences on low back pain, disability, and response to treatment. *Phys Ther*, 2011. 91(5): 712-21.

99. Straus SE, Richardson WS, Glasziou P, et al. *Evidence-Based Medicine: How to Practice and Teach EBM*. 3rd ed, ed. M. Parkinson. 2005, New York: Elsevier, Churchill Livingstone.

100. Boersma K, Linton SJ. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *Clin J Pain*, 2005. 21(1): 38-43; discussion 69-72.

101. van der Windt D, Hay E, Jellema P, et al. Psychosocial interventions for low back pain in primary care: lessons learned from recent trials. *Spine (Phila Pa 1976)*, 2008. 33(1): 81-9.

102. George SZ, Fritz JM, Bialosky JE, et al. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine (Phila Pa 1976)*, 2003. 28(23): 2551-60.

103. George SZ, Fritz JM, McNeil DW. Fear-avoidance beliefs as measured by the fear-avoidance beliefs questionnaire: change in fear-avoidance beliefs questionnaire is predictive of change in self-report of disability and pain intensity for patients with acute low back pain. *Clin J Pain*, 2006. 22(2): 197-203.

104. Sullivan MJ, Stanish WD. Psychologically based occupational rehabilitation: the Pain-Disability Prevention Program. *Clin J Pain*, 2003. 19(2): 97-104.

105. Burton AK, Waddell G, Tillotson KM, et al. Information and advice to patients with back pain can have a positive effect. A randomized controlled trial of a novel educational booklet in primary care. *Spine (Phila Pa 1976)*, 1999. 24(23): 2484-91.

106. Vowles KE, Gross RT. Work-related beliefs about injury and physical capability for work in individuals with chronic pain. *Pain*, 2003. 101(3): 291-8.

107. Woby SR, Watson PJ, Roach NK, et al. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? *Eur J Pain*, 2004. 8(3): 201-10.

108. Mannion AF, Junge A, Taimela S, et al. Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy. *Spine (Phila Pa 1976)*, 2001. 26(8): 920-9.

109. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain*, 1998. 14(3): 209-15.

110. Grotle M, Vollestad NK, Brox JI. Screening for yellow flags in first-time acute low back pain: reliability and validity of a Norwegian version of the Acute Low Back Pain Screening Questionnaire. *Clin J Pain*, 2006. 22(5): 458-67.

111. Fritz JM, George SZ. Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther*, 2002. 82(10): 973-83.

112. Cleland JA, Fritz JM, Brennan GP. Predictive validity of initial fear avoidance beliefs in patients with low back pain receiving physical therapy: is the FABQ a useful screening tool for identifying patients at risk for a poor recovery? *Eur Spine J*, 2008. 17(1): 70-9.

113. de Souza FS, Marinho Cda S, Siqueira FB, et al. Psychometric testing confirms that the Brazilian-Portuguese adaptations, the original versions of the Fear-Avoidance Beliefs Questionnaire, and the Tampa Scale of Kinesiophobia have similar measurement properties*. Spine (Phila Pa 1976)*, 2008. 33(9): 1028-33.

114. Woby SR, Watson PJ, Roach NK, et al. Adjustment to chronic low back pain--the relative influence of fear-avoidance beliefs, catastrophizing, and appraisals of control*. Behav Res Ther*, 2004. 42(7): 761-74.

115. Lundberg M, Frennered K, Hagg O, et al. The impact of fear-avoidance model variables on disability in patients with specific or nonspecific chronic low back pain*. Spine (Phila Pa 1976)*, 2011. 36(19): 1547-53.

116. Foster NE, Thomas E, Bishop A, et al. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care*. Pain*, 2010. 148(3): 398-406.

117. Nicholas MK. Obstacles to recovery after an episode of low back pain; the 'usual suspects' are not always guilty*. Pain*, 2010. 148(3): 363-4.

118. Hazard RG, Haugh LD, Reid S, et al. Early prediction of chronic disability after occupational low back injury*. Spine (Phila Pa 1976)*, 1996. 21(8): 945-51.

119. Linton SJ, Hallden K. *Risk factors and the natural course of acute and recurrent musculoskeletal pain: developing a screening instrument.* in *Proceedings of the 8th World Congress on Pain*. 1997. Seattle: IASP Press.

120. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment*. Arthritis Rheum*, 2008. 59(5): 632-41.

121. Fritz JM, Beneciuk JM, George SZ. Relationship Between Categorization With the STarT Back Screening Tool and Prognosis for People Receiving Physical Therapy for Low Back Pain*. Phys Ther*, 2011. 91(5): 722-32.

122. Kongsted A, Johannesen E, Leboeuf-Yde C. Feasibility of the STarT back screening tool in chiropractic clinics: a cross-sectional study of patients with low back pain*. Chiropr Man Therap*, 2011. 19: 10.

123. Boersma K, Linton SJ. Psychological processes underlying the development of a chronic pain problem: a prospective study of the relationship between profiles of psychological variables in the fear-avoidance model and disability. *Clin J Pain*, 2006. 22(2): 160-6.

124. Hill JC, Dunn KM, Main CJ, et al. Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. *Eur J Pain*, 2010. 14(1): 83-9.

125. Swartz RJ, Schwartz C, Basch E, et al. The king's foot of patient-reported outcomes: current practices and new developments for the measurement of change. *Qual Life Res*, 2011. 20(8): 1159-1167.

126. Pincus T, Santos R, Breen A, et al. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum*, 2008. 59(1): 14-24.

127. Hay EM, Dunn KM, Hill JC, et al. A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. The STarT Back Trial Study Protocol. *BMC Musculoskelet Disord*, 2008. 9: 58.

128. Morso L, Albert H, Kent P, et al. Translation and discriminative validation of the STarT Back Screening Tool into Danish. *Eur Spine J*, 2011.

129. Gusi N, del Pozo-Cruz B, Olivares PR, et al. The Spanish version of the "STarT Back Screening Tool" (SBST) in different subgroups. *Aten Primaria*, 2011. 43(7): 356-61.

130. Waddell G, Newton M, Henderson I, et al. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, 1993. 52(2): 157-68.

131. Jacob T, Baras M, Zeev A, et al. Low back pain: reliability of a set of pain measurement tools. *Arch Phys Med Rehabil*, 2001. 82(6): 735-42.

132. Pfingsten M, Kroner-Herwig B, Leibing E, et al. Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). *Eur J Pain*, 2000. 4(3): 259-66.

133. Staerkle R, Mannion AF, Elfering A, et al. Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. *Eur Spine J*, 2004. 13(4): 332-40.

134. Kovacs FM, Muriel A, Medina JM, et al. Psychometric characteristics of the Spanish version of the FAB questionnaire. *Spine (Phila Pa 1976)*, 2006. 31(1): 104-10.

135. Swinkels-Meewisse IE, Roelofs J, Verbeek AL, et al. Fear-avoidance beliefs, disability, and participation in workers and non-workers with acute low back pain. *Clin J Pain*, 2006. 22(1): 45-54.

136. George SZ, Fritz JM, Childs JD. Investigation of elevated fear-avoidance beliefs for patients with low back pain: a secondary analysis involving patients enrolled in physical therapy clinical trials. *J Orthop Sports Phys Ther*, 2008. 38(2): 50-8.

137. Sullivan M, Bishop S, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychological Assessment*, 1995. 7(4): 524-532.

138. Van Damme S, Crombez G, Bijttebier P, et al. A confirmatory factor analysis of the Pain Catastrophizing Scale: invariant factor structure across clinical and non-clinical populations. *Pain*, 2002. 96(3): 319-24.

139. Crombez G, Eccleston C, Baeyens F, et al. When somatic information threatens, catastrophic thinking enhances attentional interference. *Pain*, 1998. 75(2-3): 187-98.

140. Osman A, Barrios FX, Gutierrez PM, et al. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. *J Behav Med*, 2000. 23(4): 351-65.

141. Woby SR, Roach NK, Urmston M, et al. Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. *Pain*, 2005. 117(1-2): 137-44.

142. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*, 2001. 16(9): 606-13.

143. Huang FY, Chung H, Kroenke K, et al. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med*, 2006. 21(6): 547-52.

144. Damush TM, Wu J, Bair MJ, et al. Self-management practices among primary care patients with musculoskeletal pain and depression. *J Behav Med*, 2008. 31(4): 301-7.

145. Lowe B, Kroenke K, Herzog W, et al. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord*, 2004. 81(1): 61-6.

146. Spielberger CD, Gorsuch RL, Lushene RE, et al. *Manual for the state and trait anxiety inventory (Form Y)*. 1983, Palo Alto, CA: Consulting Psychologists Press.

147. Bialosky JE, Bishop MD, Robinson ME, et al. Spinal manipulative therapy has an immediate effect on thermal pain sensitivity in people with low back pain: a randomized controlled trial. *Phys Ther*, 2009. 89(12): 1292-303.

148. Bieling PJ, Antony MM, Swinson RP. The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behav Res Ther*, 1998. 36(7-8): 777-88.

149. Kabacoff RI, Segal DL, Hersen M, et al. Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *J Anxiety Disord*, 1997. 11(1): 33-47.

150. Kennedy BL, Schwab JJ, Morris RL, et al. Assessment of state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatr Q*, 2001. 72(3): 263-76.

151. Jensen MP, Turner JA, Romano JM, et al. Comparative reliability and validity of chronic pain intensity measures. *Pain*, 1999. 83(2): 157-62.

152. Bolton JE. Accuracy of recall of usual pain intensity in back pain patients. *Pain*, 1999. 83(3): 533-9.

153. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine*, 2005. 30(11): 1331-4.

154. Hudson-Cook N, Tomes-Nicholson K, Breen A, *A revised Oswestry disability questionnaire*, in *Back Pain: New Approaches to Rehabilitation and Education*, M.O. Roland and J.R. Jenner, Editors. 1989, Manchester University Press: New York, NY. p. 187-204.

155. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther*, 2001. 81(2): 776-88.

156. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)*, 2000. 25(22): 2940-52; discussion 2952.

157. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)*, 2000. 25(24): 3115-24.

158. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*, 1983. 8(2): 141-4.

159. Waddell G, Somerville D, Henderson I, et al. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine (Phila Pa 1976)*, 1992. 17(6): 617-28.

160. Fritz JM, Piva SR. Physical impairment index: reliability, validity, and responsiveness in patients with acute low back pain. *Spine (Phila Pa 1976)*, 2003. 28(11): 1189-94.

161. Kleinbaum DG, Kupper LL, Muller KE. *Applied regression analysis and other multivariable methods*. 3rd ed. 1998, Pacific Grove, CA: Brooks/Cole Publishing Company.

162. Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. *Ann Intern Med*, 1993. 118(3): 201-10.

163. George SZ, Delitto A. Clinical examination variables discriminate among treatment-based classification groups: a study of construct validity in patients with acute low back pain. *Phys Ther*, 2005. 85(4): 306-14.

164. Milligan GA, Cooper MC. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 1985. 50(2): 159-179.

165. Ketchen DJ, Shook CL. The application of cluster analysis in strategic management research: an analysis and critique. *Strategic Management Journal*, 1996. 17: 441-458.

166. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 5th ed. 2007, Boston, MA: Pearson /Allyn and Bacon.

167. Meyers LW, Gamst G, Guarino AJ. *Applied Multivariate Research: Design and Interpretation*. 2006, Thousand Oaks, CA: Sage Publications.

168. Field A. *Discovering Statistics Using SPSS.* 2nd ed. 2005, Thousand Oaks, CA: SAGE Publications.

169. Hill JC, Vohora K, Dunn KM, et al. Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *Clin J Pain*, 2010. 26(9): 783-7.

170. Main CJ, Foster N, Buchbinder R. How important are back pain beliefs and expectations for satisfactory recovery from back pain? *Best Pract Res Clin Rheumatol*, 2010. 24(2): 205-17.

171. Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*, 2007. 147(7): 478-91.

172. Beneciuk JM, Bishop MD, George SZ. Clinical prediction rules for physical therapy interventions: a systematic review. *Phys Ther*, 2009. 89(2): 114-24.

173. Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. *J Orthop Sports Phys Ther*, 2007. 37(6): 290-302.

174. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*, 2011.

175. Smeets RJ, Vlaeyen JW, Kester AD, et al. Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *J Pain*, 2006. 7(4): 261-71.

176. Leeuw M, Goossens ME, van Breukelen GJ, et al. Exposure in vivo versus operant graded activity in chronic low back pain patients: results of a randomized controlled trial. *Pain*, 2008. 138(1): 192-207.

177. Kroenke K, Wu J, Bair MJ, et al. Reciprocal relationship between pain and depression: a 12-month longitudinal analysis in primary care. *J Pain*, 2011. 12(9): 964-73.

BIOGRAPHICAL SKETCH

Jason Beneciuk, PT, DPT, MPH, FAAOMPT earned his Master of Public Health (MPH) with an epidemiology concentration from the University of Florida in 2011. He was recognized as a Fellow of the American Academy of Orthopaedic Manual Physical Therapists (AAOMPT) after completing his advanced manual therapy training at the University of St. Augustine for Health Sciences in 2006. He earned his Doctor of Physical Therapy (DPT) degree from the University of St. Augustine for Health Sciences in 2002 and his Bachelor of Science (BS) degree in biology from the Richard Stockton College of New Jersey in 1996. Jason is a member of various professional associations including the American Physical Therapy Association, the Florida Physical Therapy Association, the American Academy of Orthopaedic Manual Physical Therapists, and the American Pain Society.